
W H I T E P A P E R

Agentic Orchestration for ANZ Professional Services

Trust, Guardrails, and Progressive Autonomy in Accounting, Tax, Audit, and Compliance Advisory

A Dual-Zone Architecture for Deploying Autonomous Sub-Agents and Governed Decision Agents Across Regulated Workflows in New Zealand and Australia

Published by Zato Origin New Zealand Limited

Authors: Saurav J Bansal, Dr Srinivas Kishan Anapu, Sanjeev Reddy Bora

March 2026 | Version 2.0

Classification: Public

Executive Summary

The deployment of AI agents in the accounting, tax, audit, and compliance advisory professions across New Zealand and Australia demands more than technical capability; it demands trust. As agentic AI systems evolve from passive assistants into orchestrated multi-agent architectures capable of autonomous task execution, the central challenge shifts from what agents can do to which tasks they should perform autonomously, which tasks require governed human oversight, and how the boundary between these zones is defined, enforced, and audited.

This white paper introduces ZATO's framework for Agentic Orchestration in ANZ Professional Services, built around a Dual-Zone Architecture that separates agent work into two clearly defined operational domains. The Autonomous Execution Zone encompasses preparatory, analytical, and classification tasks—such as data gathering, document normalisation, transaction classification, workpaper drafting, and anomaly detection—where sub-agents operate with speed and efficiency under lightweight oversight. The Governed Decision Zone encompasses judgment-dependent, legally consequential, and opinion-forming activities where progressive autonomy, execution guardrails, risk-tiered human-in-the-loop controls, and full reasoning traceability ensure that professional standards and regulatory obligations are met.

A four-stage autonomy model—Shadow, Assisted, Supervised, and Constrained—governs agent progression within the Governed Decision Zone, ensuring that agent independence is earned through demonstrated performance rather than assumed by default. A template-driven Job Orchestrator sequences agent workflows across both zones, managing the handover from autonomous sub-agents to governed decision agents with precision. A comprehensive Trust and Governance framework, including prompt injection defence and data exfiltration prevention, underpins the entire architecture.

Positioned within the emerging discourse on responsible and trustworthy AI agents—particularly as articulated by Gartner, NIST, and the OECD—this paper provides a structured reference for professional services firms, regulators, and technology leaders seeking to deploy agentic AI systems that are not merely powerful, but demonstrably safe, auditable, and aligned with the specific obligations of accounting, tax, audit, and advisory practice in the ANZ region.

1. The Trust Imperative in Agentic AI for Professional Services

1.1 From Chatbots to Autonomous Agents

The evolution of AI in professional services has traversed three distinct phases. The first generation comprised rule-based systems and robotic process automation (RPA), executing predefined scripts without contextual understanding. The second generation introduced large language models (LLMs) as conversational assistants, capable of answering questions and generating text but lacking the ability to act on behalf of users. The third generation—now emerging—deploys agentic AI systems that can plan, reason, execute multi-step workflows, and autonomously interact with external tools and data sources.

Source: Gartner, “Multiagent Systems in Enterprise AI,” December 2025; IBM Think, “AI Agents in 2025: Expectations vs. Reality,” November 2025

This transition from passive intelligence to active agency introduces a fundamentally different risk profile. When an AI agent can prepare a tax filing, trigger a compliance review, or generate an audit workpaper, the question is not whether it can perform the task, but whether it should perform the task without human governance—and if so, which specific components of the task are safe for autonomous execution and which require human judgment.

1.2 Why Trust Cannot Be Assumed

In accounting, tax, audit, and compliance advisory, trust is not a feature to be marketed—it is a precondition for deployment. Professional services firms operate under statutory obligations, professional standards (ISA, ASA, NZ IFRS, AASB), and fiduciary duties that require demonstrable diligence in every action taken on a client’s behalf. An AI agent operating within these environments must meet the same evidentiary burden as a human professional.

Research from the OECD’s AI Policy Observatory and NIST’s AI Risk Management Framework (AI RMF 1.0) emphasises that trustworthy AI systems must be transparent, accountable, secure, and subject to human oversight—principles that are especially critical when AI systems move from advisory roles to execution roles.

Source: OECD AI Policy Observatory, “Trustworthy AI Principles,” 2024; NIST AI 100-1, “AI Risk Management Framework,” January 2023

1.3 The Regulatory Landscape in ANZ Professional Services

In New Zealand, Inland Revenue’s intensified enforcement posture, backed by NZ\$116 million in additional funding, and the forthcoming provisions of the Taxation (Annual Rates for 2025–26) Bill demand that any AI-assisted compliance output be traceable, verifiable, and defensible. In Australia, new climate reporting assurance requirements and expanded compliance programmes compound this burden. Firms deploying AI agents for tax compliance, audit support, or advisory engagement must demonstrate that every agent action is governed, auditable, and subject to appropriate human oversight proportionate to the risk and regulatory sensitivity of the specific task.

Source: CPA Australia, “Well-funded IRD drives an aggressive compliance approach,” 2025; AASB Climate Reporting Standards 2025

2. The Dual-Zone Architecture: Autonomous Sub-Agents and Governed Decision Agents

A critical architectural decision in deploying agentic AI within professional services is determining where autonomous execution is safe and where governed human oversight is essential. ZATO addresses this through a Dual-Zone Architecture that classifies all agent work into two operationally distinct domains, each with its own governance model, oversight requirements, and agent types.

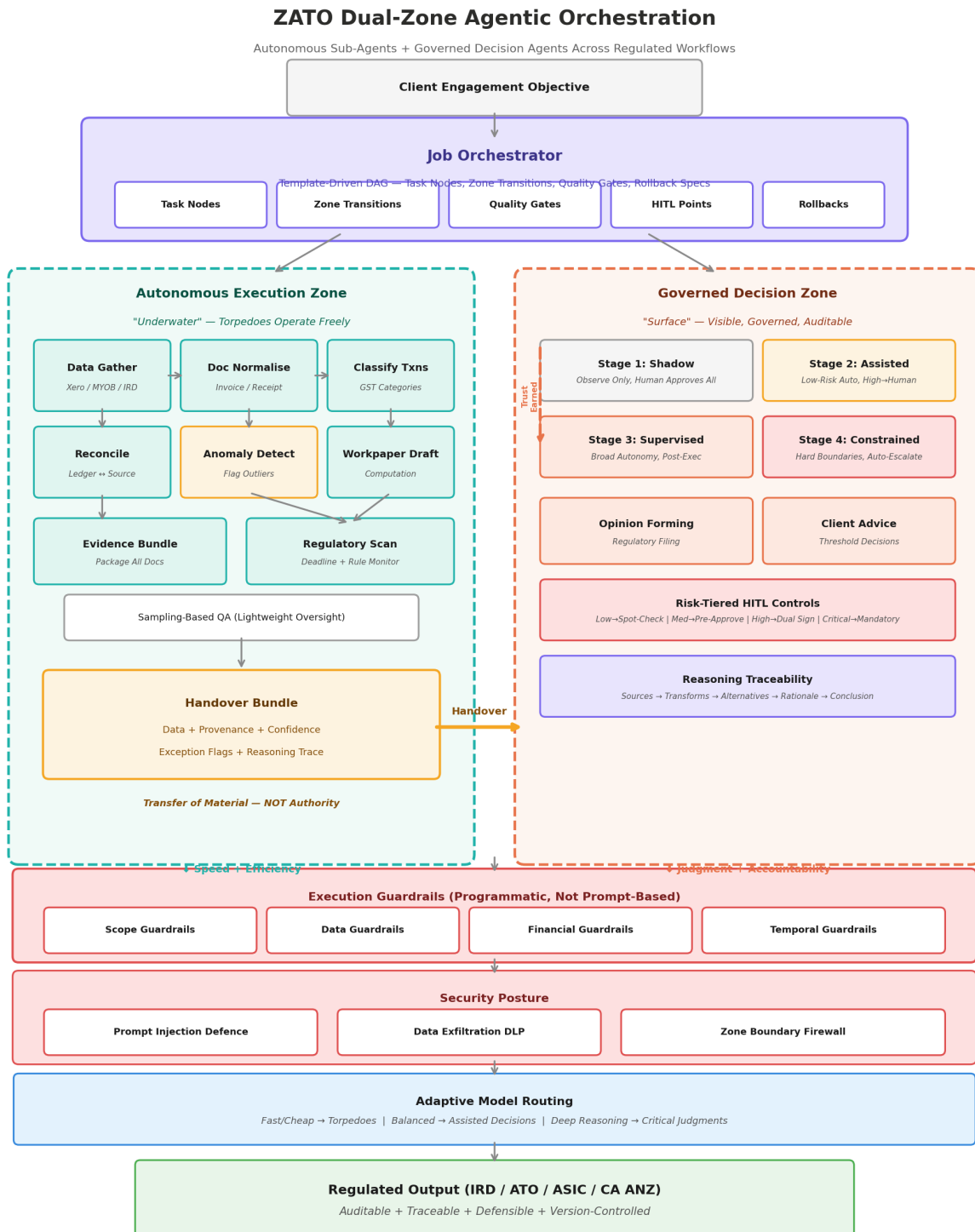


Figure 1: ZATO Dual-Zone Master Architecture

2.1 The Classification Principle

Not all tasks in accounting, tax, audit, and compliance carry equal regulatory risk or require equal professional judgment. Preparing a bank reconciliation is fundamentally different from forming an audit opinion. Normalising a set of invoices for GST classification is fundamentally different from advising on the tax consequences of a cross-border restructuring. The Dual-Zone Architecture makes this distinction explicit and enforceable.

The classification principle operates on a simple test: does the task require professional judgment, form a legally consequential opinion, or produce an output that will be communicated externally to a client, regulator, or third party without further human review? If the answer is no, the task falls within the Autonomous Execution Zone. If the answer is yes, the task falls within the Governed Decision Zone.

2.2 The Autonomous Execution Zone

The Autonomous Execution Zone encompasses preparatory, analytical, and classification tasks that sub-agents—specialised, lightweight agents optimised for speed and throughput—can perform with minimal human oversight. These are tasks in which the agent gathers, organises, classifies, or prepares material for subsequent human or governed-agent review. The outputs of these tasks feed into the Governed Decision Zone but do not, by themselves, constitute professional advice, regulatory submissions, or client-facing opinions.

ZATO defines eleven categories of regulated-safe autonomous work within the ANZ professional services context:

Table 1: Autonomous Execution Zone — Regulated-Safe Task Categories

#	Task Category	Description
1	Gathering Data	Collecting and ingesting structured and unstructured data from connected source systems and API endpoints
2	Normalising Documents	Standardising document formats, extracting key fields, and structuring unstructured content into consistent schemas
3	Reconciling Records	Matching transactions across systems and identifying discrepancies for review against source ledgers
4	Classifying Transactions	Categorising financial transactions against chart of accounts, regulatory taxonomies, or supply type classifications
5	Drafting Workpapers	Preparing structured analytical workpapers with supporting schedules, cross-references, and computation details
6	Preparing Checklists	Generating regulatory and procedural checklists with completion status tracking and gap identification
7	Proposing Exceptions	Identifying items that fall outside normal parameters and flagging them for human evaluation and disposition
8	Detecting Anomalies	Applying statistical and pattern-based analysis to identify outliers, irregularities, and unusual transaction patterns
9	Routing Cases	Assessing incoming work items and directing them to appropriate specialist agents or human reviewers based on domain and risk
10	Preparing First-Pass Narratives	Generating draft analytical commentary and explanatory narratives for human refinement and professional review
11	Generating Evidence Bundles	Assembling and packaging supporting documentation into structured bundles for human review and sign-off

These eleven categories share a common characteristic: they produce intermediate outputs that inform professional judgment but do not themselves constitute that judgment. A sub-agent that classifies a transaction as zero-rated for GST purposes proposes the classification; a human practitioner (or a governed decision agent operating under HITL controls) confirms it. A sub-agent that drafts an audit workpaper prepares a document; the engagement team reviews, adjusts, and approves it.

2.3 The Governed Decision Zone

The Governed Decision Zone encompasses all tasks that involve professional judgment, opinion formation, regulatory submission, or client-facing communication. These tasks are performed by governed decision agents operating under the progressive autonomy model (Section 3), subject to execution guardrails, risk-tiered HITL controls, and full traceability of reasoning.

Examples of Governed Decision Zone activities in ANZ professional services include:

- **Opinion-forming:** Audit opinions, tax position conclusions, compliance assessments, advisory recommendations.
- **Regulatory submissions:** Tax return filing, GST/BAS lodgement, suspicious activity reports, and statutory filings.
- **Client communications:** Advice letters, engagement deliverables, management reports, board presentations.
- **Judgment-dependent analyses:** Transfer pricing determinations, materiality assessments, going-concern evaluations, restructuring advice.
- **Threshold decisions:** Whether to escalate a matter, qualify an opinion, disclose a risk, or recommend a course of action.

2.4 The Handover Mechanism: From Sub-Agent to Decision Agent

The boundary between the Autonomous Execution Zone and the Governed Decision Zone is enforced through a structured handover protocol managed by the Job Orchestrator. When a sub-agent completes its autonomous task, it packages its outputs into a standardised handover bundle that includes the processed data or document, a provenance record detailing sources and applied transformations, a confidence score for proposed classifications or judgments, and any flags or exceptions identified during processing.

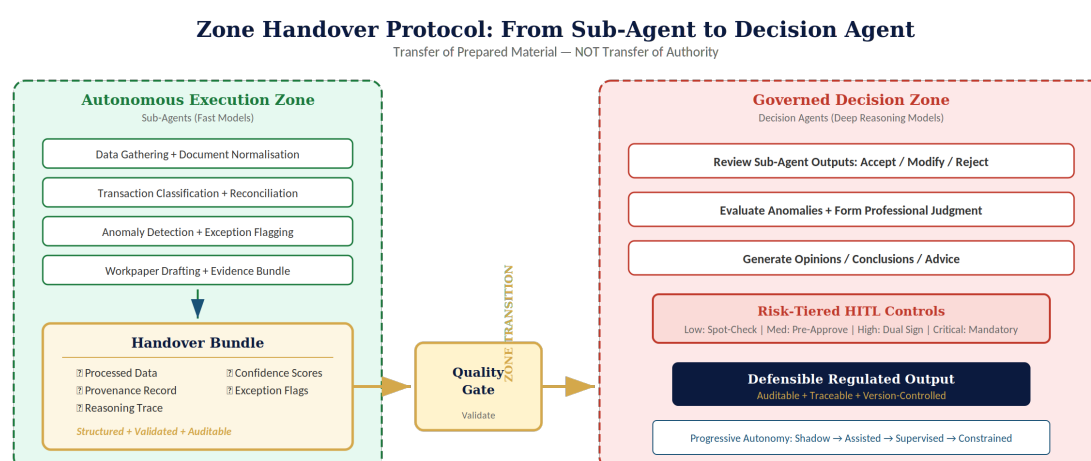


Figure 2: Zone Handover Protocol — From Sub-Agent to Decision Agent

The governed decision agent receives this handover bundle as its input, along with the full audit trail of the sub-agent's work. Critically, the decision agent is not bound by the sub-agent's proposed classifications or drafts—it reviews them as a senior practitioner would review a junior's work, accepting, modifying, or rejecting as professional judgment requires. The handover is a transfer of prepared material, not a transfer of authority.

Source: Gartner, "Multiagent Systems in Enterprise AI," December 2025; ISACA, "No Looking Back: Transforming Audit with Artificial Intelligence," June 2025

3. Progressive Autonomy in the Governed Decision Zone

Within the Governed Decision Zone, ZATO implements a four-stage maturity model that progressively scales agent autonomy as trust is earned through demonstrated performance, compliance, and governance maturity. This model applies specifically to decision agents—agents performing judgment-dependent, opinion-forming, or legally consequential work—not to the sub-agents operating within the Autonomous Execution Zone, which are governed by the lighter oversight model described in Section 2.2.

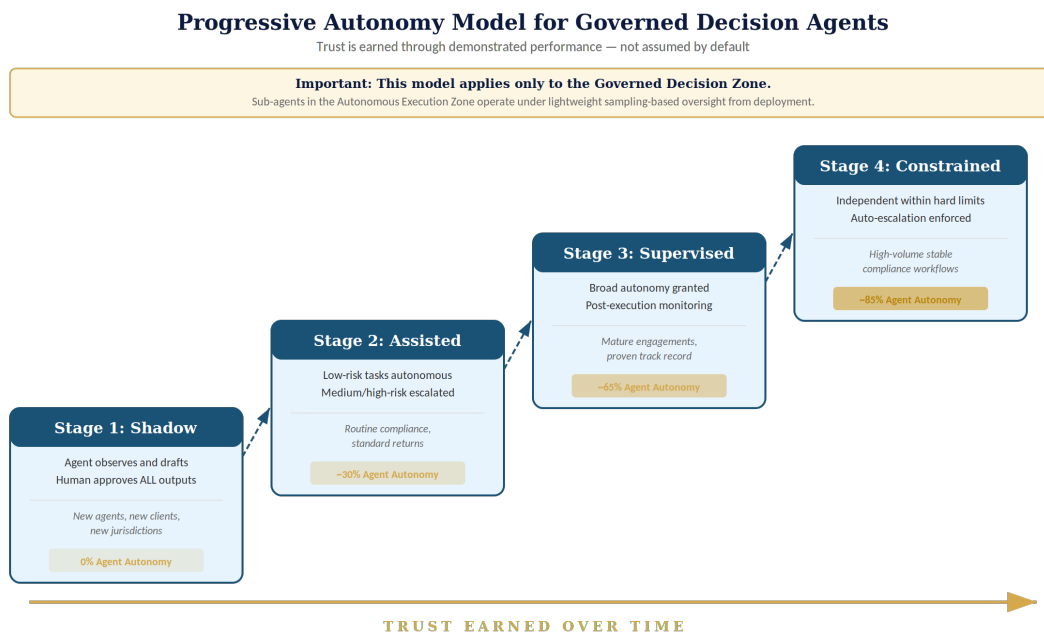


Figure 3: Four-Stage Progressive Autonomy Model for Governed Decision Agents

3.1 Stage 1: Shadow Mode

In Shadow Mode, decision agents observe and analyse but do not act on governed tasks. The agent processes incoming handover bundles from sub-agents, generates recommendations, and produces draft outputs, but every deliverable is presented to a human professional for review before any external action is taken. The agent's outputs are logged alongside the human's final decisions, creating a comparative record that enables performance benchmarking over time.

- **Agent capability:** Full analytical processing of sub-agent handover bundles, recommendation generation, draft output creation.
- **Human role:** Mandatory review and approval of every output before execution or external communication.
- **Trust mechanism:** Performance comparison between agent recommendations and human decisions, building a statistical trust baseline.
- **Applicable contexts:** Initial deployment of any new decision agent type; first engagement with a new client; entry into a new regulatory jurisdiction.

3.2 Stage 2: Assisted Mode

In Assisted Mode, the decision agent begins to execute lower-risk governed tasks autonomously while routing higher-risk or ambiguous tasks for human approval. The boundary between autonomous execution

and human review is determined by a risk classification engine that evaluates each task against predefined criteria including regulatory sensitivity, financial materiality, client risk profile, and task complexity.

- **Agent capability:** Autonomous execution of lower-risk governed tasks; flagging and routing of medium and high-risk tasks for human review.
- **Human role:** Approval required for tasks above the defined risk threshold; periodic spot-checking of autonomous executions.
- **Trust mechanism:** Risk-tiered escalation policies; error rate monitoring; client feedback integration.
- **Applicable contexts:** Routine compliance conclusions; standard tax return sign-off components; well-precedented advisory responses.

3.3 Stage 3: Supervised Mode

In Supervised Mode, the decision agent operates with broad autonomy across most governed task categories, with human oversight shifting from pre-approval to post-execution review. The human professional monitors agent performance through dashboards, exception reports, and audit trails rather than reviewing every individual output.

- **Agent capability:** Broad autonomous execution across standard and moderate-complexity governed tasks; self-flagging of edge cases.
- **Human role:** Post-execution monitoring; exception-based intervention; periodic quality assurance reviews.
- **Trust mechanism:** Statistical process control; anomaly detection; client outcome tracking; regulatory compliance audits.
- **Applicable contexts:** Mature client engagements with established agent performance history; well-understood regulatory domains.

3.4 Stage 4: Constrained Autonomous Mode

Constrained Autonomous Mode represents the highest level of decision agent independence within ZATO's framework—but critically, it is not unconstrained autonomy. The agent operates independently within explicitly defined boundaries, with hard guardrails that prevent execution outside its sanctioned scope. Actions that exceed defined thresholds are automatically escalated regardless of the agent's confidence level.

- **Agent capability:** Independent operation within defined boundaries; automatic escalation when boundaries are approached or exceeded.
- **Human role:** Governance oversight; boundary definition and adjustment; escalation handling; strategic decision-making.
- **Trust mechanism:** Hard execution boundaries; automatic escalation triggers; comprehensive audit trails; regulatory compliance certification.
- **Applicable contexts:** High-volume, well-defined compliance workflows with established performance baselines and stable regulatory requirements.

Table 2: Progressive Autonomy Stages for Governed Decision Agents

Stage	Agent Capability	Human Role	Risk Scope	Trust Basis
Shadow	Observe, analyse, recommend from sub-agent outputs	Review and approve all outputs	No autonomous action	Performance benchmarking
Assisted	Execute lower-risk governed tasks autonomously	Approve medium/high-risk tasks	Low-risk governed tasks	Risk-tiered escalation
Supervised	Broad autonomous execution of governed tasks	Post-execution monitoring and exception review	Standard and moderate	Statistical process control
Constrained	Independent within hard boundaries	Governance, boundary setting, escalation	All within hard limits	Boundary enforcement

4. The Job Orchestrator: Template-Driven Agent Sequencing Across Zones

4.1 From Ad Hoc Prompts to Structured Workflows

A common failure mode in enterprise AI deployments is the reliance on ad hoc prompting—where each user interaction is treated as an independent event without structured workflow context. In accounting, tax, and audit engagements, this approach is fundamentally inadequate. Professional workflows follow defined methodologies with specific sequencing requirements, quality gates, and documentation obligations.

ZATO's Job Orchestrator addresses this by introducing template-driven agent sequencing: predefined workflow templates that specify the exact sequence of agent actions across both the Autonomous Execution Zone and the Governed Decision Zone, including the handover points between them.

4.2 Template Architecture

Each Job Template encapsulates a complete workflow as a directed acyclic graph (DAG) of agent tasks. Templates are defined declaratively and include the following components:

- **Task Nodes:** Individual agent actions with specified inputs, outputs, and success criteria. Each node is explicitly tagged as belonging to the Autonomous Execution Zone or the Governed Decision Zone.
- **Zone Transitions:** Explicit handover points where sub-agent outputs are packaged and transferred to governed decision agents, including the validation checks that must pass before handover is accepted.
- **Dependency Edges:** Sequencing constraints that prevent downstream tasks from executing before upstream prerequisites are complete.
- **Quality Gates:** Validation checkpoints that verify output integrity, regulatory compliance, and data consistency before the workflow progresses.
- **HITL Insertion Points:** Predefined positions within the Governed Decision Zone where human review and approval are mandatory, determined by the risk classification of the task and the current autonomy stage.
- **Rollback Specifications:** Defined procedures for reversing completed steps when downstream failures or quality gate rejections require workflow correction.

4.3 Workflow Example: GST Return Preparation

The following diagram illustrates a complete GST return preparation workflow as a DAG, showing how sub-agent task nodes in the Autonomous Execution Zone feed through a quality gate into decision agent nodes in the Governed Decision Zone, with rollback paths and mandatory HITL checkpoints.

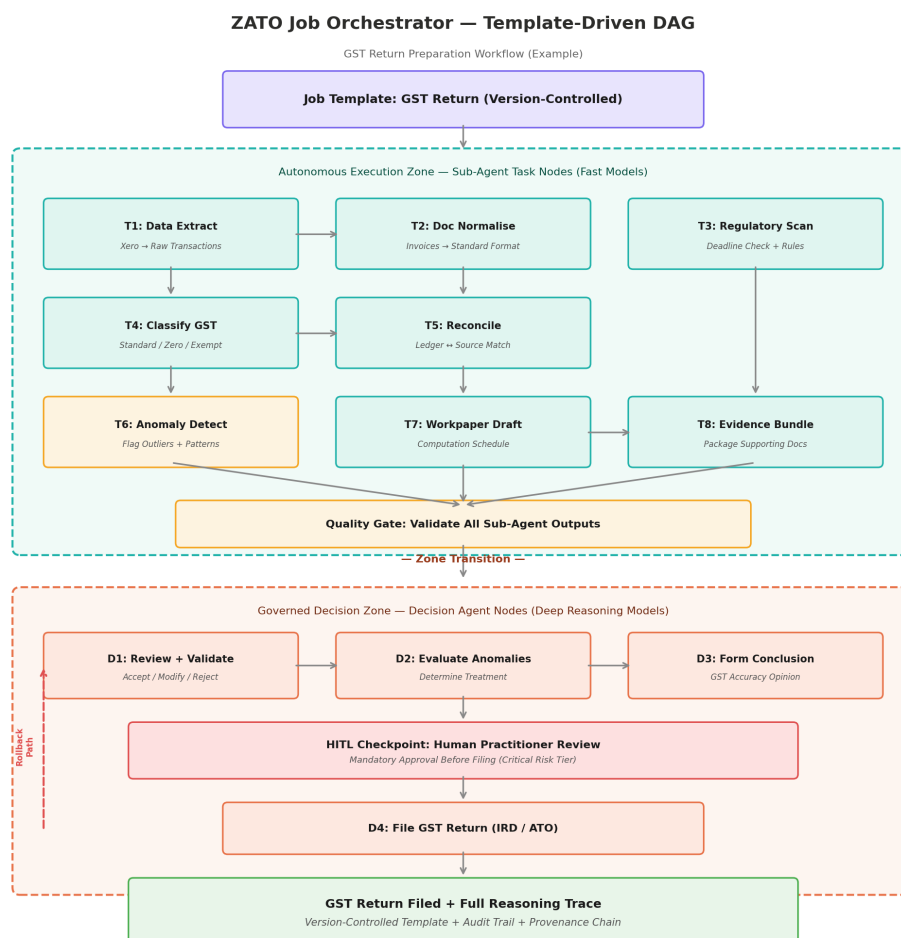


Figure 4: Job Orchestrator DAG — GST Return Preparation Workflow

4.4 Adaptive Model Routing Within Templates

Building on ZATO’s adaptive model routing capability, the Job Orchestrator assigns different AI models to different task nodes based on the cognitive demands of each step. Sub-agents in the Autonomous Execution Zone are powered by fast, cost-efficient models optimised for throughput. Decision agents in the Governed Decision Zone engage balanced-reasoning or deep-reasoning models calibrated to the complexity of the governed task.

This mirrors the staffing model of professional services firms: data extraction and reconciliation are handled by capable but efficient team members, while complex advisory judgments engage senior specialists.

4.5 Template Governance and Version Control

Job Templates are subject to formal governance processes including version control, change approval workflows, and audit trail documentation. When regulatory requirements change, templates are updated through a controlled process that ensures all active workflows comply with current legislative requirements. Historical template versions are retained for audit purposes.

Source: Gartner, “Emerging Practices in AI Governance for Regulated Industries,” 2025; ISACA, “No Looking Back: Transforming Audit with Artificial Intelligence,” June 2025

5. Trust and Governance Framework

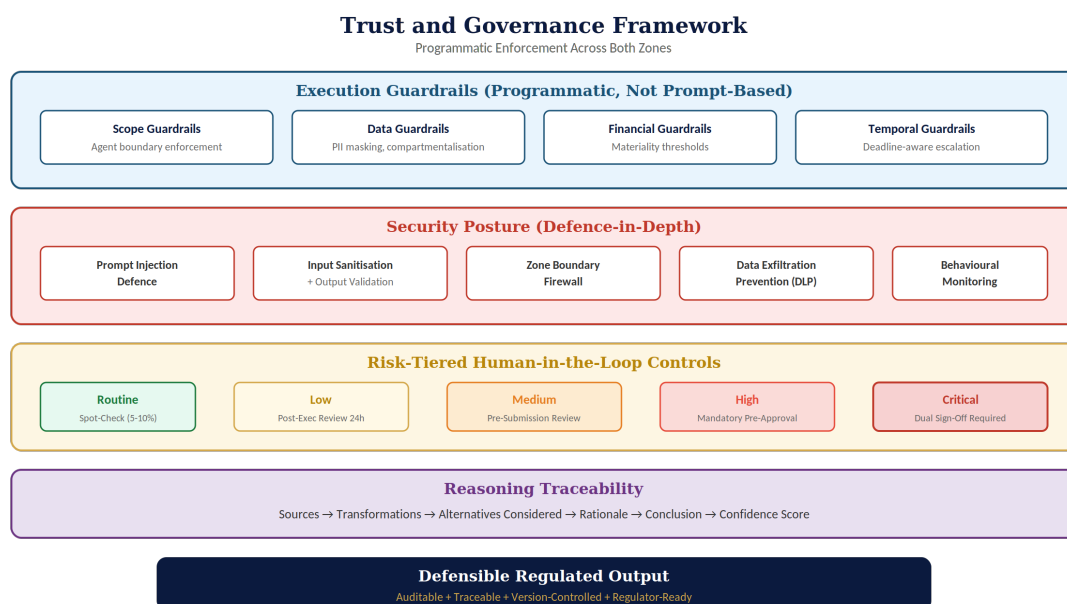


Figure 5: Trust and Governance Framework — Layered Enforcement

5.1 Execution Guardrails

Execution guardrails are hard constraints that govern agent behaviour at runtime across both zones. Unlike guidelines or recommendations, guardrails are enforced programmatically and cannot be overridden by the agent's own reasoning. ZATO implements guardrails across four dimensions:

- **Scope Guardrails:** Define the boundaries of what each agent type is permitted to do. A sub-agent tasked with transaction classification cannot autonomously escalate to forming a tax opinion. A decision agent tasked with GST compliance cannot autonomously initiate a transfer pricing analysis.
- **Data Guardrails:** Restrict the data sources an agent can access and the data it can output. Client confidential information is compartmentalised, preventing cross-client data contamination. PII is subject to masking and access controls aligned with the Privacy Act 2020 (NZ) and the Privacy Act 1988 (AU).
- **Financial Guardrails:** Impose materiality thresholds that trigger mandatory human review. Any agent-generated output involving financial figures above defined thresholds requires human validation before external communication or filing.
- **Temporal Guardrails:** Enforce deadline awareness and prevent agents from executing time-sensitive actions outside approved windows. Filing deadline proximity triggers escalation protocols that increase human oversight.

5.2 Risk-Tiered Human-in-the-Loop (HITL) Controls

ZATO's HITL framework implements a risk-tiered model where the intensity and nature of human involvement is calibrated to the risk profile of each specific task. HITL controls apply primarily within the Governed Decision Zone, while the Autonomous Execution Zone operates under sampling-based quality assurance.

Table 3: Risk-Tiered HITL Control Matrix

Risk Tier	Zone	Example Tasks	HITL Requirement	Escalation Path
Routine	Autonomous Execution	Data gathering, normalisation, reconciliation, classification	Post-execution sampling (5–10%)	Team lead
Low	Governed Decision	Standard compliance conclusions, routine advisory	Post-execution review within 24h	Senior practitioner
Medium	Governed Decision	Tax return preparation, compliance assessments	Pre-submission review	Manager/Senior Manager
High	Governed Decision	Advisory opinions, multi-jurisdictional analysis	Mandatory pre-approval	Partner/Director
Critical	Governed Decision	Audit sign-off, regulatory submissions, formal advice	Dual review with sign-off	Partner + compliance

5.3 Reasoning Traceability

Every agent action within ZATO—in both the Autonomous Execution Zone and the Governed Decision Zone—generates a structured reasoning trace that documents the agent's analytical process from input to output.

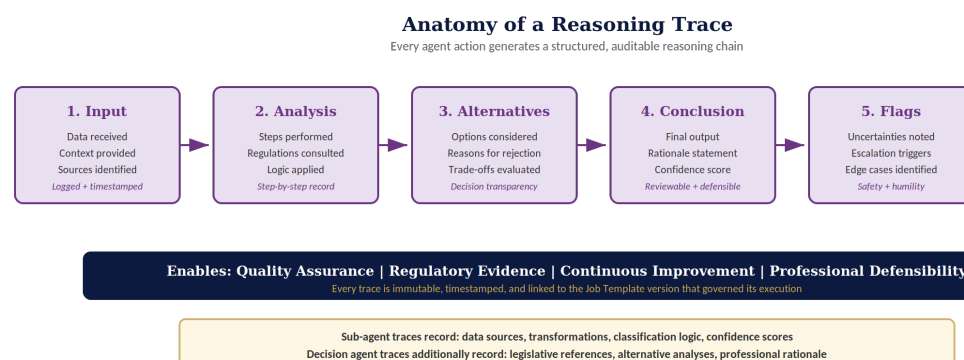


Figure 6: Anatomy of a Reasoning Trace

For sub-agents, this trace records data sources accessed, transformations applied, classification logic, and confidence scores. For decision agents, the trace additionally records the legislative and regulatory references consulted, alternative analyses considered and reasons for rejection, and the rationale for the final professional conclusion.

Reasoning traces serve multiple governance functions: quality assurance verification, regulatory compliance evidence, continuous improvement data, and—critically—the ability to demonstrate to regulators and professional body reviewers exactly how an AI-assisted professional output was produced.

Source: NIST AI 100-1, “AI Risk Management Framework,” January 2023; OECD, “Advanced AI Assistants: Governance Challenges,” 2024

6. Prompt Injection Defence and Security Posture

6.1 The Prompt Injection Threat

Prompt injection represents one of the most significant security risks in agentic AI systems. Unlike traditional software vulnerabilities that exploit code flaws, prompt injection exploits the natural language interface through which agents receive instructions. An attacker can embed malicious instructions within seemingly benign data—a client document, an email, or a data feed—that cause the agent to deviate from its intended behaviour, exfiltrate data, or execute unauthorised actions.

In professional services, where agents process confidential client information and generate legally consequential outputs, the threat is especially acute. The Dual-Zone Architecture provides an inherent defence layer: sub-agents in the Autonomous Execution Zone have limited scope and cannot escalate to governed actions, meaning a compromised sub-agent cannot autonomously file a tax return or issue professional advice.

Source: OWASP, "Top 10 for Large Language Model Applications," 2024

6.2 ZATO's Defence-in-Depth Approach

ZATO implements a multi-layered defence against prompt injection and related adversarial attacks:

- **Input Sanitisation:** All external inputs are processed through a sanitisation layer that identifies and neutralises potential injection patterns before they reach the agent's reasoning engine.
- **Instruction Hierarchy:** ZATO enforces a strict hierarchy of instruction authority. System-level instructions (defined by the Job Template and governance framework) cannot be overridden by user-level or data-level inputs.
- **Zone Boundary Enforcement:** Sub-agents cannot invoke decision agent capabilities or access governed action endpoints, regardless of instructions received. The zone boundary acts as a hard architectural firewall.
- **Output Validation:** Agent outputs are validated against expected patterns, regulatory constraints, and historical baselines before execution.
- **Behavioural Monitoring:** Runtime monitoring tracks agent behaviour for anomalous patterns that may indicate successful injection.
- **Isolation and Sandboxing:** Each agent operates within a sandboxed environment with defined resource access boundaries.

6.3 Data Exfiltration Prevention

ZATO addresses the broader risk of data exfiltration through comprehensive data loss prevention (DLP) controls. Agent outputs are screened for sensitive data patterns (tax identification numbers, bank account details, client identifiers) before any external communication. Cross-client data isolation ensures that information from one client engagement cannot leak into another through shared agent context or model state.

7. Positioning Within the Responsible AI Discourse

7.1 Alignment with International Frameworks

ZATO's Dual-Zone Architecture and progressive autonomy framework are deliberately aligned with the principal international frameworks for trustworthy and responsible AI:

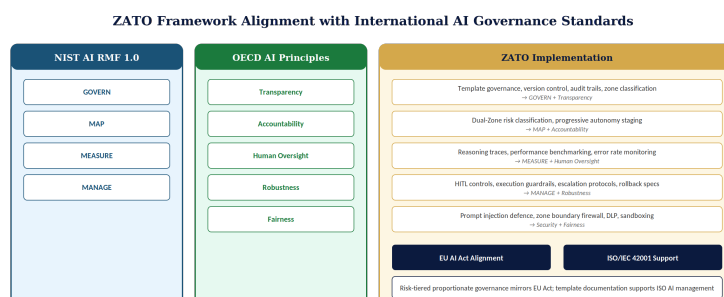


Figure 7: ZATO Framework Alignment with International AI Governance Standards

- **NIST AI Risk Management Framework (AI RMF 1.0):** ZATO's progressive autonomy model maps to NIST's GOVERN, MAP, MEASURE, and MANAGE functions. The Dual-Zone classification implements NIST's principle of proportionate risk management.
- **OECD AI Principles:** The framework's emphasis on transparency (reasoning traces), accountability (audit trails), human oversight (HITL controls), and robustness (execution guardrails) directly implements the OECD's five principles for trustworthy AI.
- **EU AI Act Risk Classification:** ZATO's risk-tiered approach mirrors the EU AI Act's proportionate regulation model, positioning firms for compliance should similar frameworks emerge in the ANZ region.
- **ISO/IEC 42001 (AI Management System):** ZATO's template governance, version control, and audit trail capabilities support the emerging AI management system standard.

7.2 The Gartner Perspective on Agentic AI Governance

Gartner's research on multiagent systems emphasises that organisations must establish governance frameworks before scaling agentic deployments. The firm's 2025 guidance specifically highlights the need for explicit agent boundary definition, human override capabilities, audit trail completeness, and progressive trust calibration—all core design principles of ZATO's framework.

The Dual-Zone Architecture directly addresses Gartner's observation that agentic AI governance must distinguish between tasks that can be safely automated and tasks that require governed oversight—a distinction that many early agentic deployments fail to make explicitly.

Source: Gartner, "Multiagent Systems in Enterprise AI: Efficiency, Innovation and Vendor Advantage," December 2025

7.3 Implications for ANZ Regulators and Professional Bodies

As AI agents become prevalent in accounting, tax, and audit practice, ANZ regulators and professional bodies will need frameworks for evaluating AI-assisted professional outputs. ZATO's architecture facilitates this by providing clear delineation between autonomous preparatory work and governed professional judgment, full reasoning traceability for regulatory review, and template-based methodology documentation that enables regulators to assess the governance framework itself, not just individual outputs.

8. Applications in ANZ Professional Services

8.1 Tax Compliance and Advisory

Tax engagements exemplify the Dual-Zone Architecture in practice. Sub-agents in the Autonomous Execution Zone handle data gathering from connected source systems, transaction classification, income and deduction schedule preparation, and first-pass computation of tax positions. Decision agents in the Governed Decision Zone review these outputs, form conclusions on complex positions (transfer pricing, R&D credits, international tax), and prepare client-facing advice—all under the progressive autonomy model appropriate to the engagement's maturity and risk profile.

8.2 Audit and Assurance

Audit engagements benefit significantly from the zone separation. Sub-agents perform data analytics, workpaper preparation, reconciliation, anomaly detection, and evidence bundle assembly—tasks that consume significant junior staff time. Decision agents, operating in Shadow or Assisted Mode for new engagements, review these outputs and support the engagement team in forming conclusions on risk assessment, materiality, and audit findings. Opinion-forming activities remain at Critical-tier HITL controls regardless of autonomy stage.

8.3 Regulatory Compliance and AML/CFT

AML/CFT compliance demonstrates the importance of the handover mechanism. Sub-agents perform customer due diligence data gathering, transaction monitoring, and suspicious pattern detection. Decision agents evaluate flagged items, determine whether reporting thresholds are met, and prepare suspicious activity reports—with mandatory human sign-off before any regulatory submission.

8.4 Strategic Advisory

Complex advisory engagements—entity restructuring, succession planning, cross-border expansion—involve the highest stakes and the greatest need for professional judgment. Sub-agents provide research support, data analysis, and scenario modelling. Decision agents operate primarily in Shadow Mode, providing structured analytical support that enhances the advisory team's capacity without substituting for the professional judgment that clients engage the firm to provide.

9. Future Directions

The Dual-Zone agentic orchestration framework described in this paper represents ZATO's current architectural foundation. Several future development trajectories are anticipated:

- **Dynamic Zone Reclassification:** As agent capabilities mature and governance frameworks evolve, certain tasks currently in the Governed Decision Zone may be reclassified to the Autonomous Execution Zone, with appropriate safeguards and regulatory approval.
- **Inter-Agent Deliberation Protocols:** Future iterations will enable decision agents to engage in structured debate before synthesis, with adversarial reasoning processes that strengthen output quality.
- **Federated Trust Scoring:** Development of standardised trust metrics enabling agent performance and governance maturity to be assessed consistently across firms and jurisdictions.
- **Regulatory Sandbox Integration:** Collaboration with ANZ regulators to establish sandbox environments where agentic AI systems can be tested against real regulatory scenarios.
- **Cross-Framework Interoperability:** As the "Internet of Agents" matures, ZATO's governance framework will evolve to support secure, governed interactions between agents operating within different organisational contexts.
- **Continuous Assurance Models:** Evolution from periodic audit cycles to continuous assurance models where sub-agents monitor compliance in real time and decision agents evaluate emerging issues as they arise.

Source: Gartner, "Multiagent Systems in Enterprise AI," December 2025; NIST, "AI RMF Playbook," 2023

10. Conclusion

The deployment of agentic AI in accounting, tax, audit, and compliance advisory is not a question of if, but of how. The firms that will lead this transformation are not those that deploy agents fastest, but those that deploy agents most responsibly—with architectures that clearly delineate where autonomous execution is safe and where governed human oversight is essential.

ZATO's Dual-Zone Architecture provides this clarity. The Autonomous Execution Zone unleashes the speed and efficiency of sub-agents across eleven categories of regulated-safe preparatory work—data gathering, document normalisation, transaction classification, workpaper drafting, anomaly detection, and more—while the Governed Decision Zone ensures that professional judgment, opinion formation, and legally consequential actions remain subject to progressive autonomy controls, execution guardrails, risk-tiered human oversight, and full reasoning traceability.

The four-stage progressive autonomy model ensures that decision agent independence is earned, not assumed. The template-driven Job Orchestrator ensures that workflows across both zones are repeatable, auditable, and aligned with professional methodology. The Trust and Governance framework ensures that every agent action is constrained, traceable, and proportionately governed. And the security posture ensures that zone boundaries, data isolation, and prompt injection defences are architecturally enforced.

For the accounting, tax, audit, and advisory firms that serve New Zealand and Australian businesses, ZATO offers a path to agentic AI adoption that is not merely technically capable but professionally defensible—a critical distinction in a profession where trust is the ultimate currency.

References

1. Gartner, "Multiagent Systems in Enterprise AI: Efficiency, Innovation and Vendor Advantage," December 2025.
2. NIST, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," AI 100-1, January 2023.
3. OECD AI Policy Observatory, "Trustworthy AI Principles and Recommendations," 2024.
4. OECD, "Advanced AI Assistants: Governance Challenges and Policy Directions," 2024.
5. OWASP, "Top 10 for Large Language Model Applications," Version 1.1, 2024.
6. ISO/IEC 42001:2023, "Information Technology — Artificial Intelligence — Management System."
7. CPA Australia, "Well-funded IRD drives an aggressive compliance approach," 2025.
8. PwC New Zealand, "Tax Tips: Year in Review 2025."
9. NZ Inland Revenue, Taxation (Annual Rates for 2025–26, Compliance Simplification, and Remedial Measures) Bill Commentary, August 2025.
10. AASB, Climate Reporting and Assurance Standards, 2025.
11. ISACA, "No Looking Back: Transforming Audit with Artificial Intelligence," June 2025.
12. IBM Think, "AI Agents in 2025: Expectations vs. Reality," November 2025.
13. CPA.com/AICPA, "2025 AI in Accounting Report," June 2025.
14. Chartered Accountants Australia and New Zealand (CA ANZ), "Submission on the 2025 Occupation Shortage List Stakeholder Survey," 2025.
15. Bloomberg Tax, "Big Four Firms Roll Out AI That Can Handle Routine Tasks Solo," March 2025.
16. BDO, Audit Innovation Survey 2025.
17. Mordor Intelligence, "AI in Accounting Market Report," 2025.
18. Accountancy Age, "Big Four lag behind smaller firms in AI adoption, says ex-EY chair," August 2025.
19. EU Artificial Intelligence Act, Regulation (EU) 2024/1689, August 2024.

Agentic Orchestration for ANZ Professional Services

Trust, Guardrails, and Progressive Autonomy

Built by Zato | Governed by Design

© 2026 Zato Origin New Zealand Limited. All rights reserved.