

AI Agent–Driven Autonomous Service Delivery in Saudi Arabia: A Framework for Intelligent Telecom and Enterprise Operations under Vision 2030

WAMIQ RAFI SYED
[0009-0004-7079-4931](https://doi.org/10.64388/IREV10I4-1723620)

Abstract- The telecommunications and enterprise sectors in Saudi Arabia are progressing from digitally based service operations towards more autonomous operating models in which artificial intelligence is able to understand objectives, identify situations, coordinate tools, and carry out specific actions. This report looks at the way in which AI agents can help to achieve autonomous service delivery in both telecom and enterprise operations while at the same time meeting the requirements relating to reliability, governance, cybersecurity, and human accountability that are linked to Saudi Vision 2030. The study makes use of a structured integrative review of peer-reviewed literature published between 2020 and 2025, with a focus on agentic AI, zero-touch network and service management, intent-based networking, AIOps, service operations, AI governance, human oversight, and Saudi digital transformation. The analysis highlights five interdependent capabilities: context-aware intent interpretation, multi-source observability, agentic reasoning and orchestration, closed-loop execution, and governance-by-design. It is shown that autonomy is most believable when it is implemented as bounded, observable, and reversible automation rather than as unrestricted machine control. Research in the telecom field has developed well-established foundations in the areas of intent life cycles, assurance, closed loops, and zero-touch management, whereas research relating to enterprise services brings in aspects such as incident intelligence, workflow enhancement, and service quality. The review introduces an Agentic Autonomous Service Delivery Framework that combines operational agents with policy constraints, confidence thresholds, digital audit trails, human escalation, and continuous assurance. This framework provides a practical research programme for Saudi organisations that wish to establish resilient, scalable, and trustworthy AI-enabled service operations in line with Vision 2030.

I. INTRODUCTION

Telecommunications and the operations of business technology have got to a stage of scale and

interdependence where traditional methods of service management are being put under increasing pressure. The operational states resulting from cloud-native platforms, software-defined networks, distributed applications, edge computing, 5G services, cybersecurity measures, data platforms, and multi-vendor infrastructure change so quickly that manual teams can no longer consistently keep up with them. That is why research into zero-touch management views automation as a complete end-to-end management issue rather than as a set of separate scripts [1,2]. Likewise, intent-based networking moves the focus away from commands at the device level and instead centres on the desired outcomes, which can then be translated, activated, monitored, and assured via closed loops [3,4].

The range of what is possible in this kind of transition has been broadened by artificial intelligence. Although machine learning has for a long time enabled anomaly detection, prediction, classification, and optimization, it is only with the development of large language models and autonomous agents that planning, tool use, memory, reasoning, and natural-language interaction have become available. Thanks to these capabilities it is now conceivable to have service-delivery agents which take a business or technical objective, examine the relevant operational evidence, prepare a plan, invoke monitoring and orchestration tools, suggest or carry out changes, check the results, and pass the issue on when their level of confidence is not sufficient. Recent research into autonomous agents refers to this change as a move from passive model inference to goal-directed systems that integrate reasoning with actions in an external environment [5,6]. In the area of network management, intent systems based on large language models and experiments using agent-based fault tolerance already show how natural-language

objectives can be turned into operational actions or into coordinated diagnostic procedures [7,8].

The situation is especially pertinent to Saudi Arabia, since Vision 2030 has sped up the process of digital transformation in the areas of government, telecommunications, finance, industry and large-scale infrastructure. Saudi research on service economies regards digital transformation, service lifecycle management, smart infrastructure, data analytics and autonomous systems as interconnected elements of future service design [9]. Empirical studies of the telecommunications sector in the country also link digital transformation practices to improved operational and marketing efficiency [10], and more general economic research connects digitalisation with diversification and growth [11]. Owing to these developments, there is a strong need for operating models which are able to maintain high service quality without requiring proportionate increases in manual coordination.

Autonomous service delivery cannot simply be regarded as an obvious extension of chatbot automation. An agent which is capable of modifying a network policy, restarting a service, altering a cloud configuration, revoking access, rerouting traffic, or triggering a business workflow has what might be described as operational agency. Errors can spread throughout interconnected systems, and a locally correct action might still breach a security policy, a service-level agreement, a regulatory obligation, or a broader optimization goal. Research into AI-enabled network automation therefore places particular emphasis on trust, coordination, explainability, and performance safeguards in multi-vendor environments [12]. Likewise, governance research contends that AI systems need lifecycle controls, accountability, assurance, interpretability, and clear human oversight instead of relying just on principles [13-16].

The literature is therefore well developed in a number of areas but is scattered among different research communities. Studies in the field of telecommunications offer rigorous models covering autonomous network management, intent translation, zero-touch assurance, and closed-loop control. The literature on AIOps focuses on anomaly detection, incident management, root-cause analysis, and

operational intelligence. Research into service operations shows how AI can be used to automate routine tasks while at the same time enhancing expert judgement [17]. Research on agentic AI includes planning and tool coordination, while governance research makes clear the limits of delegation and the need for human control. However, there is still less developed work on an integrated perspective regarding service delivery that links these various elements to the Saudi context and is able to distinguish between useful autonomy and unsafe automation.

The review fills this gap by combining the latest research into a clear framework for AI agent-driven autonomous service delivery in the Saudi telecommunications and enterprise sectors. Instead of suggesting that agents take the place of operational teams, the paper looks at how limited autonomy can be used to redistribute tasks: machines carry out the high-frequency aspects of sensing, correlation, diagnosis, orchestration, and verification, while humans keep the authority for decisions that are ambiguous, of high impact, strategic, or ethically significant. This approach is in line with the evidence that AI generates value in professional services when automation and augmentation are combined rather than seen as substitutes [17].

1.1 Aim, Objectives and Research Questions

This review aims at creating a framework based on evidence to explain how AI agents can enable trustworthy autonomous service delivery in the field of telecommunications and enterprise operations in Saudi Arabia under Vision 2030.

The study has four aims. The first of these is to combine the operational capabilities needed for agent-driven service delivery in both network and enterprise environments. The second aim is to assess how intent-based management, AIOps, autonomous agents, and service assurance can be incorporated into a closed-loop operating model. The third aim is to identify the governance, cybersecurity, accountability, and human-control requirements which place constraints on safe operational autonomy. The fourth aim is to convert the synthesis into an implementation framework and a research programme appropriate for organisations in Saudi Arabia.

The review is based on three research questions: the first is to identify which technical and organizational capabilities are consistently linked with autonomous service delivery in the literature from 2020 to 2025; the second is to examine how AI agents can coordinate observation, reasoning, action, and assurance without compromising service reliability or human accountability; and the third is to determine which implementation principles are most appropriate for telecom and enterprise operations in Saudi Arabia as they aim to align autonomous service management with Vision 2030.

II. REVIEW METHODOLOGY

The paper employs a structured integrative review approach since the research problem spans a number of fields which employ different terminology and have different methodological traditions. It would be inappropriate to carry out a narrowly focused meta-analysis in order to capture architecture papers, network-management surveys, service-operations studies, governance research, or agentic-AI frameworks in a comparable statistical way. The aim is therefore analytical integration, which involves identifying the recurring mechanisms, tensions, and design requirements across the relevant evidence while still keeping the selection criteria transparent.

The time frame for the review extended from January 2020 to December 2025 in order to include the shift from AI-assisted network automation to current agentic and generative-AI methods. The searches were based on combinations of the following terms: "AI agent", "autonomous agent", "AIOps", "autonomous service delivery", "zero-touch network", "network and service management", "intent-based networking", "intent-driven network", "LLM network management", "incident management", "service operations", "AI governance", "human oversight", "telecommunications", "Saudi Arabia", and "Vision 2030". When establishing the evidence base, preference was given to peer-reviewed journal articles and high-quality conference research that had identifiable digital object identifiers. The foundational standards concepts were understood on the basis of peer-reviewed studies and were not restated as separate normative documents.

To be included, a study had to have been published between 2020 and 2025, have to be substantially relevant to at least one stage in the process of autonomous service delivery, and had to contain enough methodological or conceptual detail to allow for synthesis. Those studies which focused exclusively on algorithm performance and had no operational-management implications were excluded. In the same way, articles that talked about AI in general customer service were only included if they provided insights into service quality, trust, or the interaction between humans and agents that were relevant to operational delivery. Papers that were specific to Saudi Arabia were included provided they offered evidence regarding digital transformation, service economies, the efficiency of telecommunications, or the national technology environment. The final group of references consists of thirty items that have been chosen to ensure a balance among telecom automation, autonomous agents, AIOps and incident intelligence, service operations, governance, trust, and Saudi digital transformation.

The synthesis involved four stages of analysis. In the first, each study was assigned a code according to its operational function: intent interpretation, observation, anomaly detection, diagnosis, planning, orchestration, action, verification, governance, or human escalation. In the second, the studies were coded according to their level of autonomy, with the categories being decision support, recommended action, supervised execution, and bounded closed-loop execution. Third, the evidence from the telecom and enterprise environments was compared in order to find mechanisms that could be transferred. Fourth, the mechanisms that recur were then combined into framework layers and control principles.

The assessment focused on relevance, transparency, methodological coherence, and evidentiary fit rather than relying on a single numerical score. In order to identify well-established areas and recurring research gaps, survey and review papers were employed; empirical and prototype studies were used to back up specific claims regarding capabilities; governance papers were employed to establish the requirements for accountability and control. There were no made-up screening figures included. The method also does not regard publication frequency as evidence of

effectiveness. When the body of literature consists mainly of conceptual models or prototypes, the synthesis clearly treats those findings as indicating design evidence rather than as proof of performance at a production scale.

The main limitation is that the body of evidence is still uneven: while network automation is relatively well

developed, general-purpose agentic AI used in enterprise operations is advancing quickly and therefore has fewer long-term production studies. It is this limitation that leads to the cautious and bounded-autonomy approach in the proposed framework.

III. THEMATIC SYNTHESIS

Table 1. Evidence synthesis for AI agent-driven autonomous service delivery

Theme	Core evidence	Operational contribution	Primary risk/control
Intent and zero-touch management	Outcome-oriented intent life cycles and closed-loop assurance [1-4,29]	Translates service goals into executable, verifiable operations	Ambiguity and policy conflict; structured validation
Agentic reasoning and orchestration	Goal-directed planning, tool use and adaptive task sequencing [5-8]	Coordinates diagnosis, planning and action across heterogeneous tools	Hallucination or unsafe planning; grounding and scoped tools
AIOps and service intelligence	Alert correlation, anomaly detection and incident lifecycle intelligence [18,23,30]	Builds evidence for diagnosis and prioritization	Telemetry gaps and false correlation; multi-source evidence
Human control and governance	Lifecycle oversight, accountability, trust and institutional controls [13-16,19,24-26]	Defines permissions, escalation, assurance and auditability	Automation bias and weak oversight; risk-tiered authority
Saudi digital transformation	Service-economy autonomy, telecom efficiency and digitalization [9-11]	Links autonomous operations to scalable national digital services	Localization, data and institutional constraints; contextual adaptation

3.1 From Workflow Automation to Bounded Operational Agency

Traditional automation involves carrying out pre-defined procedures; agentic automation, on the other hand, includes conditional planning and the selection of tools. This difference is significant in the context of service management. While a rule can restart a process after a specific alarm has been triggered, an operational agent is able to combine several observations, deduce a probable causal chain, choose among different diagnostic tools, compare various remedies, and modify its plan when the environment changes. The agentic-AI descriptions identify autonomy, goal orientation, memory, planning, and interaction with the environment as the key features of the new paradigm [5,6]. However, the practical usefulness of such systems depends much more on careful integration with reliable tools and on having

explicit permissions than on anthropomorphic 'intelligence'.

In enterprise service operations, AI has proven its value by reducing repetitive diagnostic or administrative tasks and at the same time extending the capacity of experts [17]. Research into intelligent alert and incident management also indicates that operational intelligence involves the entire process from detection through to correlation, prioritisation, diagnosis and resolution rather than merely comprising a single classification task [18]. This lends support to an agent model in which various tools or dedicated agents carry out specific functions while a coordinator keeps track of the context throughout the incident lifecycle. In terms of continuity, the information collected during detection can be used to aid diagnosis, remediation, verification,

documentation and post-incident learning without the need for manually reconstructing it each time.

Autonomy should, however, be stepped up in degree. Research into human control indicates that supervision is not enough when the human is given an unclear recommendation too late to be able to take effective action [14,19]. A more robust design regards the roles of human and machine as part of a coordinated system; low-risk and reversible actions can be delegated according to preestablished policies, medium-risk

actions may require confirmation or dual approval, and high-impact actions can stay advisory. In this way, graduated authority turns the idea of having a human in the loop from a mere slogan into a measurable control model which is based on the criticality of the action, uncertainty, reversibility, and blast radius.

3.2 Intent as the Interface Between Business Outcomes and Machine Action

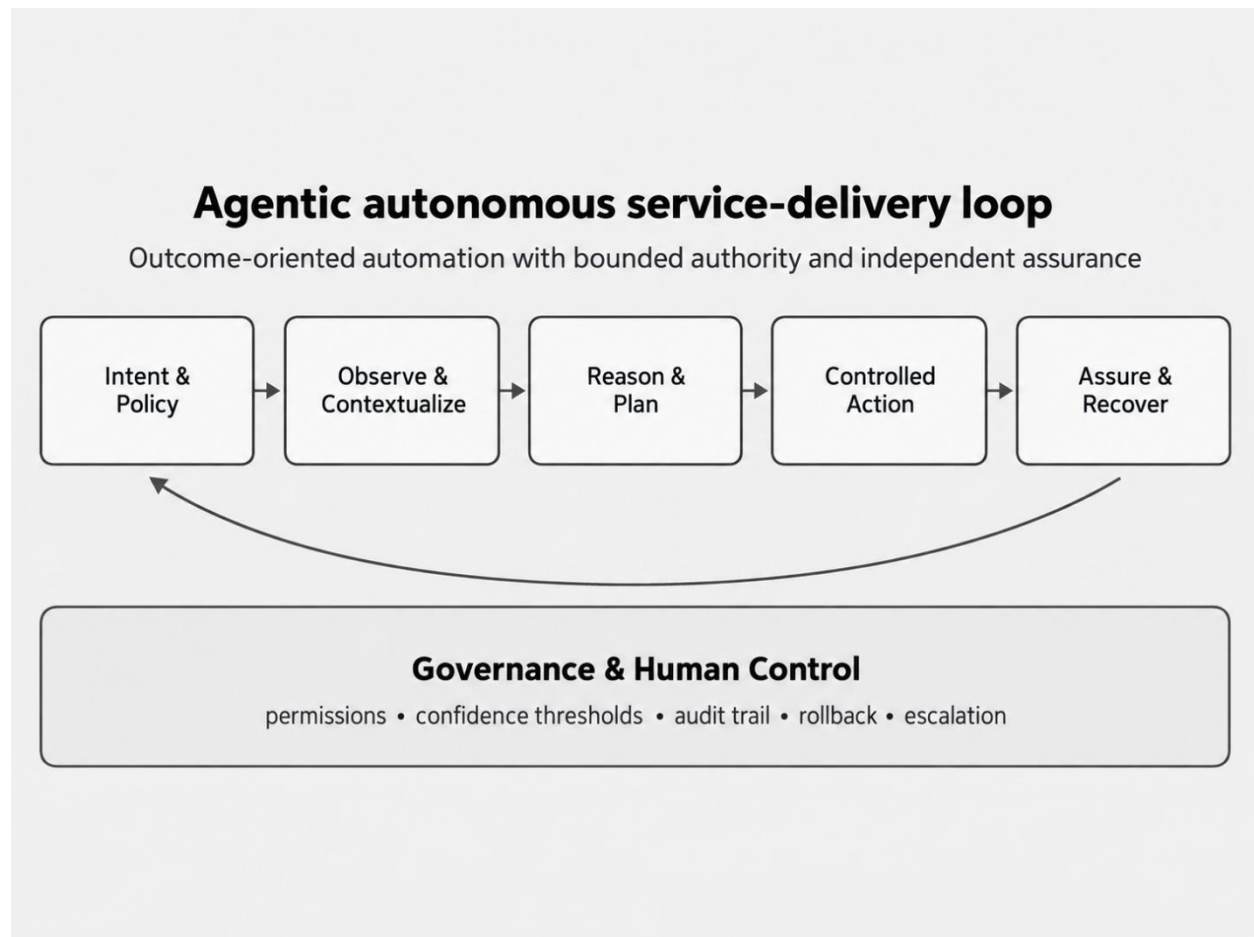


Figure 1. Agentic autonomous service-delivery loop with governance spanning the operational cycle.

Agent-driven service delivery has one of the most clear-cut bases in telecommunications research: intent-based networking. Rather than asking administrators to define device configurations, intent systems record the desired outcomes and then convert them into policies, configurations, and assurance objectives. Surveys have identified intent translation, conflict resolution, activation, monitoring, and

assurance as the key functions throughout the lifecycle [3,4]. Zero-touch management applies this principle to closed-loop operation in which the current state of the network is constantly checked against the desired outcomes [1,2,29].

This approach can be directly applied to enterprise operations. For example, a service-management intent

like "restore the payment service within the critical SLA without violating change-freeze controls" includes an objective, some constraints, and an assurance condition. The agent has to break down the intent into operational tasks, identify the relevant telemetry, determine the allowable actions, and check that the restoration has been achieved. The main innovation is that the objective stays fixed while the sequence of actions can adapt to the context.

Studies based on large language models show that the use of natural language can lessen the difficulty involved in acquiring and translating intent [7,20,21]. However, it would be wrong to regard language models as authoritative policy makers. Natural-language expressions may be ambiguous, contradictory, or incomplete. The safest approach is to use a two-stage process of interpretation: the agent first transforms the natural language into a structured intent schema and then a deterministic policy layer checks the scope, constraints, identities, dependencies, and allowed actions. Retrieval methods can connect the interpretation to network documentation, service catalogs, configuration models, or operational knowledge, and conflict-detection logic can be used to determine whether the new intent contradicts existing objectives [21,22].

For Saudi companies that manage complex multi-vendor estates, intent abstraction also provides a solution to a real management issue. Since business outcomes usually involve teams that are set up according to infrastructure, cloud, network, security, application, and vendor lines, an intent-based control plane is able to define common service outcomes without needing each stakeholder to handle every platform directly. In this way, it supports the Vision 2030 goal of having digitally enabled and integrated services while at the same time keeping the technical specialization beneath the abstraction [9-11].

3.3 Observability, AIOps and Causal Service Intelligence

Autonomy cannot exist unless there is reliable observation. Agents working with incomplete telemetry are able to arrive at the wrong diagnosis more quickly than a human team. AIOps research thus forms a second basis for the field: by combining logs,

metrics, traces, events, topology, configuration, and service context it is possible to detect anomalies and to explain the operational situation. As intelligent incident-management surveys point out, reducing the number of alerts by itself is not enough; useful systems have to link the symptoms to the incidents, the likely causes, the affected services, and the available remedies [18]. Research into anomaly detection within DevOps also shows how machine learning can be used to identify abnormal operational states in continuous delivery environments [23].

The approach to service delivery needs to change from a focus on infrastructure observability to one on service observability. An autonomous agent must realise not just that there has been an increase in CPU usage or that packet loss has occurred, but also which customer journey, business process, regulatory control, or contractual objective is in danger. For this reason, there should be a contextual graph that connects technical resources to services, owners, dependencies, SLAs, security classifications, and recent changes. This kind of context enables causal reasoning and helps reduce the risk of improving a single metric at the cost of overall end-to-end performance.

Multiple sources also increase confidence. For example, a possible network fault could be supported by telemetry information, changes to the topology, a deterioration in user experience, the configuration history, and correlated alarms. In the case of an enterprise incident, it would be necessary to look at application logs, cloud metrics, identity events, database performance, and recent deployment data. Agents should therefore regard operational conclusions as bundles of evidence rather than as the output of a single model. Confidence can then be based on the diversity and consistency of the evidence, not just on the probability score of one classifier.

It is especially crucial in the field of telecommunications, since self-optimizing components are capable of interacting in unforeseeable manners. Studies concerning AI-enabled multi-vendor management indicate that separate cognitive functions may aim at local objectives which conflict with one another or cause errors to spread [12]. Autonomous delivery thus

demands coordination between agents and among their optimization goals. The service-level orchestrator must resolve conflicting actions in light of global constraints, and the assurance agent should check that the overall result has in fact led to an improvement in the intended service outcome.

3.4 Closed-Loop Execution, Verification and Learning

What gives an autonomous service-delivery agent its main operational value is not the ability to explain what ought to happen, but rather its capacity to carry out a controlled loop running from detection to a verified result. This is what zero-touch research terms continuous sensing, analysis, decision, execution, and assurance [1,2]. Intent frameworks likewise stress that activation must include assurance against the original intent [3,4]. Agentic service delivery must therefore be based on reversible loops not on one-way automation.

A strong loop starts off with an evaluation of the state and an analysis of the impact. The agent first identifies the service that is affected, finds out if the problem is already known, gets the relevant runbooks and policies, makes an estimate of the likely effect of the possible actions, and then checks the relevant permissions. The execution is carried out using approved orchestration tools instead of allowing direct access without constraints. Each action is recorded together with the evidence used, the policy invoked, the predicted outcome and the actual response. Verification involves comparing the telemetry after the action has been taken with both the technical health indicators and the original service intent.

If verification fails, the agent ought not just keep making up answers. The framework must include bounded retry rules, alternative plans, rollback thresholds, and escalation procedures. This illustrates the important difference between resilience and uncontrolled persistence. An autonomous system which persistently pursues a goal can become dangerous if its understanding of the environment is incorrect. Research into governance of advanced agents points out this very issue: as capability increases so does the need for testing, containment, and institutional controls [15].

Learning should also be controlled. The results of incidents can lead to updates in retrieval knowledge, the recommended actions, the causal graphs, or the model training data, but production systems must tell the difference between immediate operational memory and validated organizational learning. Although an agent might keep the context during an incident, any changes to approved runbooks, policies, or models need to be reviewed. This kind of separation helps to maintain service stability and stops a single unusual incident from quietly changing future behaviour.

3.5 Governance, Trust and Human Accountability

The literature clearly opposes the idea that technical autonomy eliminates organizational accountability. When AI-governance reviews are carried out, they show that there are repeatedly identified needs concerning data, systems, processes, stakeholders, transparency, risk management, and oversight [13,16]. Further, research into governance affordances states that assurance and interpretability should be built into the entire system lifecycle [24]. Governance therefore has to be something that can be put into practice: it must be incorporated into identity controls, action scopes, approval rules, evidence logging, policy checks, and emergency stop mechanisms.

Trust is just as much about operational effectiveness as it is about rhetoric. Research involving AI agents that have been delegated duties indicates that people's acceptance of such agents is based on their perceived usefulness, the quality of the service provided, and a number of different trust factors [25]. More general research into trust highlights the importance of explainability, performance, fairness, and communication [26]. In an operational context, trust must be gained through consistent behaviour. The agent should state what it has observed, what it plans to do, why the action is allowed, how confident it is, what systems might be affected, how the action can be undone, and what success will appear like. While these explanations do not have to reveal the internal workings of the model, they must offer enough operational justification to enable accountable decision-making.

Human oversight also has to take into account the realistic limits of human cognition. It will fail if a

person is asked simply to approve an algorithmic decision, particularly when the reviewer does not have enough time, lacks the necessary context, or is unable to challenge the system [19]. That is why the proposed model makes use of exception-based oversight, whereby humans concentrate on cases that have a high impact, low confidence, conflicting objectives, an unusual topology, are security sensitive, or have irreversible consequences. Routine, low-risk operations are left to be carried out automatically, but the system keeps evidence at all times so that afterwards humans can reconstruct the decisions.

For Saudi telecommunications and enterprise organisations, this approach to governance by design promotes both innovation and institutional confidence. Although Vision 2030 places a demand on speed, scale and the delivery of intelligent services, the national transformation at the same time leads to a greater dependence on digital infrastructure; reliability, cybersecurity, privacy and continuity have thus become strategic requirements for digital growth rather than being constraints outside it [9-11].

IV. AGENTIC AUTONOMOUS SERVICE DELIVERY FRAMEWORK

The synthesis presents a six-layer model for AI agent-driven autonomous service delivery. In the first layer, which is Intent and Policy, the business goals, technical objectives, service-level commitments, security rules, change controls, and regulatory constraints are all expressed in a machine-readable format. Although natural-language interaction can be used to capture the intent, it is only through structured validation that it can be established whether the intent is complete, authorized, and free of conflicts.

The next level is Observability and Context, which includes telemetry, events, logs, traces, topology, configuration, tickets, knowledge, customer-experience signals, and business-service mappings. The aim here is to create a constantly updated

operational context rather than just having a dashboard full of separate indicators. The third level is Agentic Reasoning, whereby specialised agents are able to carry out correlation, diagnosis, impact analysis, planning, security validation, or knowledge retrieval, while a coordinating agent is responsible for maintaining the overall end-to-end goal. Although LLMs can be useful for synthesis and planning, high-risk decisions must be based on evidence that has been retrieved and be constrained by deterministic controls.

The fourth level is Controlled Action, which involves agents carrying out their tasks via approved APIs, automation platforms, network controllers, ITSM workflows, cloud orchestrators, and security systems. The permissions granted are based on the principle of least privilege and are limited in time and focused on the service objective. The fifth level is Assurance and Recovery, at which independent checks are carried out to verify technical health, customer impact, compliance with the SLA, compliance with policy, and the presence of any unintended side effects; if the assurance process fails, a rollback, re-planning, or escalation is initiated rather than allowing open-ended autonomous experimentation.

The sixth layer is Governance and Human Control, covering all of the other layers. It establishes the different autonomy classes, sets the confidence thresholds, determines the level of action criticality, specifies the approval requirements, outlines evidence retention procedures, maintains records of both model and prompt versions, ensures segregation of duties, guarantees auditability, and includes a means of carrying out emergency intervention. It also offers a way of slowly increasing autonomy; an action can start off as one that is merely a recommendation, move on to supervised execution once enough evidence has been gathered, and only then become bounded automatic execution when both reliability and reversibility have been demonstrated.

Table 2. Six-layer Agentic Autonomous Service Delivery Framework

Layer	Primary function	Agent responsibility	Human/control requirement
1. Intent & Policy	Define goals, constraints and service outcomes	Interpret and structure requested outcome	Authorization, conflict checks, policy ownership
2. Observability & Context	Build live operational state	Collect, correlate and contextualize evidence	Data quality, provenance, access control
3. Agentic Reasoning	Diagnose and plan	Decompose goals, compare causes and select actions	Grounding, confidence thresholds, explainable rationale
4. Controlled Action	Execute bounded change	Call approved orchestration tools and APIs	Least privilege, change controls, reversible actions
5. Assurance & Recovery	Verify outcome and recover safely	Test service state, rollback or re-plan	Independent checks, retry limits, escalation
6. Governance & Human Control	Constrain and audit the full lifecycle	Preserve evidence and trigger human intervention	Autonomy classes, audit trail, stop authority, accountable owner

In telecom operations the framework is capable of supporting fault localization, dynamic capacity adjustment, intent-driven configuration, service assurance and controlled self-healing; in enterprise operations the same approach can be used for incident

triage, cloud remediation, service-desk automation, deployment diagnostics, identity workflows and cross-platform service restoration. The common element in the design is the service outcome rather than the technology silo.

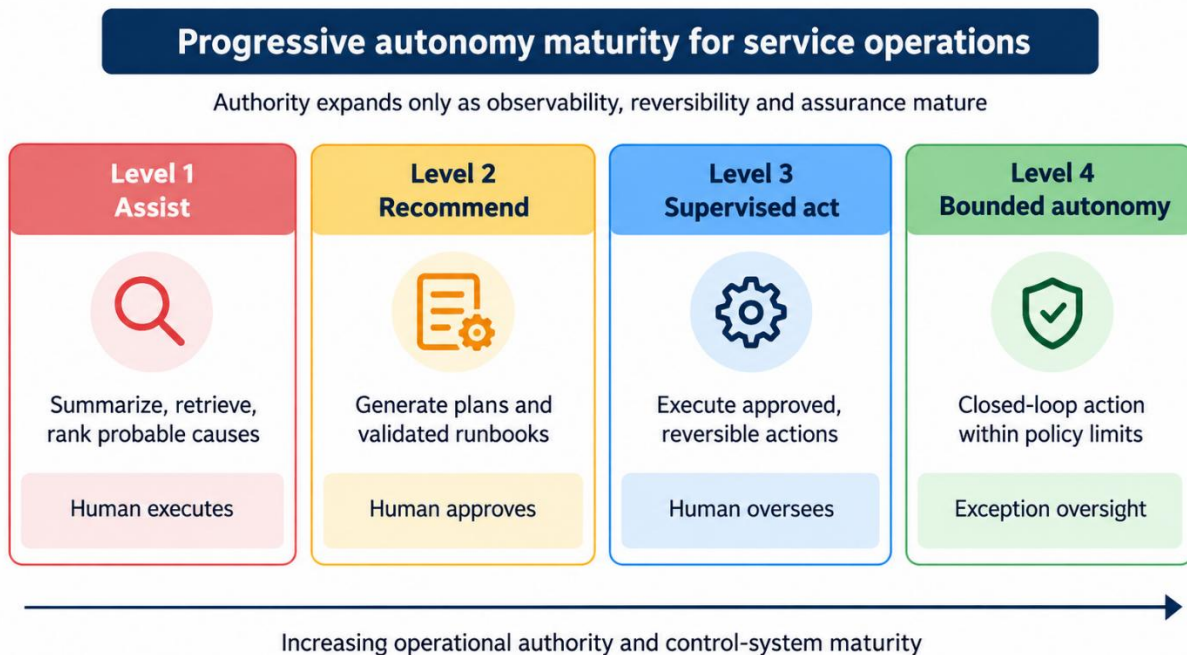


Figure 2. Progressive autonomy maturity model for telecom and enterprise service operations

V. DISCUSSION

The review shows that the best argument for AI agents in the area of service delivery is not "full autonomy" but rather coordinated autonomy. The technical literature already includes a number of the components needed for this kind of operation: intent lifecycle management [3,4], zero-touch closed loops [1,2], intelligent incident management [18], LLM-assisted intent translation [7,20-22], and autonomous fault-management prototypes [8]. What agentic AI contributes is a flexible coordination layer which can keep track of the context across a variety of different tools and adjust the sequence of tasks according to the situation.

It is for this reason that autonomous service delivery is different from conventional robotic process automation. Rigid workflows are efficient where the path is known; it is when the aim is clear but the route has to be determined through evidence gathered during the execution that agents become valuable. There are many instances of this kind in the telecom and enterprise operations sectors. A degraded service, for example, might have a number of possible causes, the available remedial actions may depend on up-to-date capacity or on the change history, and assurance may show that a technically successful action had not in fact restored the customer's outcome. Agent planning is thus most useful in the areas where deterministic tools do not apply.

The evidence also indicates that it is more important for a building to be well-designed than for the model to be novel. A model that is highly capable but which is connected directly to production systems without having been grounded in context, without having its policies validated, and without independent assurance is less useful than a less capable agent working within a well-structured control system. This finding is supported by governance research, which regards responsible AI as an organisational and lifecycle capability rather than as a feature of the model [13,16,24]. The practical consequence is that when making investments it is necessary to give equal priority to service graphs, telemetry quality, policy engineering, orchestration interfaces, identity controls, and auditability as one is given to model selection.

Another consequence is related to organizational structure. Since incidents are now more likely to span across the areas of service desk, network operations, cloud operations, security operations, and application support, AI agents tend to blur the traditional distinctions between them. The framework does not get rid of specialization; instead, it establishes a shared coordination network. Humans then take on the role of designing policies, verifying complex diagnoses, being responsible for exception handling, and supervising the quality of automation. This is similar to the automation-and-augmentation approach found in professional service operations, in which AI takes over repetitive tasks while at the same time making specialist expertise more scalable [17].

Another consequence of this is that operational autonomy should be assessed at the level of the service and not at the level of the model. Since accuracy in individual prediction tasks cannot show that an agent will maintain customer outcomes when the dependencies change, evaluation must combine technical correctness with restoration time, policy compliance, the success of a rollback, the quality of escalations, the amount of workload on operators, and the stability of the decisions when faced with repeated operational conditions.

The framework for Saudi Arabia corresponds with the general shift from digitization to intelligent service ecosystems. Research into services in Saudi Arabia stresses that digital transformation, smart infrastructure, lifecycle management and autonomous systems are all part of the future service economies and are thus closely related components [9]. Digitalization is also connected to greater economic diversification [11]. In this light, autonomous service delivery should be seen as supporting infrastructure for the country's digital objectives, since resilient operations enable cloud platforms, digital government, telecom services, smart infrastructure and AI-powered enterprises to grow with reduced operational bottlenecks.

VI. RESEARCH GAPS AND FUTURE DIRECTIONS

There are five areas that should be given priority. First, there is a need for longitudinal research that compares agent-driven operations with conventional automation in terms of reliability, mean time to restore, change failure rate, false remediation, and human workload; since prototype success is not enough to support safety-critical assertions. Second, research should focus on establishing formal methods for assessing agent confidence based on heterogeneous evidence, rather than depending on the models' self-reports. Third, multi-agent coordination must include more effective conflict-resolution mechanisms when the objectives of network performance, security, cost, and customer experience come into conflict.

Fourth, there should be benchmarks in the field of service management which assess the entire operational cycle, covering diagnosis, action, rollback, and assurance. At present, most benchmarks tend to focus merely on reasoning or anomaly detection. Fifth, research specifically concerned with Saudi Arabia should look into the questions of governance, language, data sovereignty, and the impact on the workforce arising from autonomous operations in the sectors of telecommunications, government, finance, energy, and large digital infrastructure projects.

Future designs should also investigate the use of digital twins and simulation sandboxes as pre-execution controls. Prior to carrying out a significant change, an agent could test the proposed action against a representative environment, estimate the extent of the impact, and compare the expected results. These kinds of mechanisms could serve as a practical link between systems that provide only recommendations and those with a higher degree of bounded autonomy.

VII. PRACTICAL AND POLICY IMPLICATIONS

Organisations should start by focusing on specific use cases in which the actions can be reversed, the telemetry system is well developed, and success can be measured. Appropriate initial examples are incident summarisation, ranking of the most probable cause, automated collection of evidence, recommending

runbooks, carrying out low-risk restarts, making capacity adjustments within set limits, and carrying out verification after a change. Autonomy should only be increased once monitoring has shown that performance is stable.

Management ought to establish an autonomy register which sets out the actions that agents are permitted to recommend, prepare, execute with approval, or execute automatically. For each of these categories, confidence thresholds, rollback procedures, evidence requirements, and accountable owners should be specified. In this way, governance becomes an operational practice and offers a defensible foundation for auditing autonomous actions.

VIII. LIMITATIONS

Since it brings together diverse kinds of evidence, this review does not provide statistical effect sizes for autonomous service delivery. Furthermore, as agentic AI is developing more rapidly than long-term academic evaluation, certain claims regarding its architecture are still prospective. The paper has deliberately restricted its citations to the period 2020 to 2025, a move which ensures a more up-to-date perspective but at the same time omits earlier foundational studies. Lastly, there is currently a lack of empirical evidence based in Saudi Arabia regarding autonomous operations, and it is therefore necessary to make a careful distinction between the relevance of the findings in the Saudi context and direct local validation.

IX. CONCLUSION

AI agents provide a realistic route from fragmented automation to autonomous service delivery that is focused on outcomes, provided that autonomy is constrained by context, policy, assurance, and human accountability. The relevant literature indicates that the telecommunications industry already has well-developed building blocks in the form of zero-touch management, intent-based networking, closed-loop assurance, and cognitive network control. Research into enterprise services and AIOps contributes incident intelligence, workflow augmentation, and cross-platform operational reasoning, while research

on governance identifies the institutional controls needed for safe delegation.

The framework suggested incorporates these different elements into six layers: intent and policy, observability and context, agentic reasoning, controlled action, assurance and recovery, and governance and human control. Its main idea is that an agent must not be judged solely on its ability to act; rather, it should be assessed on its ability to act within its authority, to explain the reasons for its action, to preserve evidence, to verify the resulting state of the service, to reverse any failure, and to escalate uncertainty.

For Saudi Arabia, this approach provides a practical means of linking AI innovation with the level of service reliability needed by Vision 2030. While autonomous service delivery can reduce operational friction and enhance scalability in both telecom and enterprise settings, sustainable value will ultimately rely on careful engineering rather than on maximum autonomy. The most defensible future therefore isn't one of completely human-free operations; instead, it is a gradually more autonomous operating system in which machines take care of speed and complexity whereas people continue to have meaningful control over goals, risk, and accountability.

REFERENCES

- [1] Benzaid C, Taleb T. AI-Driven Zero Touch Network and Service Management in 5G and Beyond: Challenges and Research Directions. *IEEE Network*. 2020;34(2):186-194. [IEEE](#)
- [2] Coronado E, Behraves R, Subramanya T, Fernandez-Fernandez A, Siddiqui MS, Costa-Perez X, Riggio R. Zero Touch Management: A Survey of Network Automation Solutions for 5G and 6G Networks. *IEEE Communications Surveys & Tutorials*. 2022;24(4):2535-2578. [IEEE](#)
- [3] Pang L, Yang C, Chen D, Song Y, Guizani M. A Survey on Intent-Driven Networks. *IEEE Access*. 2020;8:22862-22873. [IEEE](#)
- [4] Leivadeas A, Falkner M. A Survey on Intent-Based Networking. *IEEE Communications Surveys & Tutorials*. 2023;25(1):625-655. [IEEE](#)
- [5] Acharya DB, Kuppan K, Divya B. Agentic AI: Autonomous Intelligence for Complex Goals-A Comprehensive Survey. *IEEE Access*. 2025;13:18912-18936. [IEEE](#)
- [6] Wang L, Ma C, Feng X, Zhang Z, Yang H, Zhang J, et al. A Survey on Large Language Model Based Autonomous Agents. *Frontiers of Computer Science*. 2024;18:186345. [Springer](#)
- [7] Mekrache A, Ksentini A, Verikoukis C. Intent-Based Management of Next-Generation Networks: an LLM-Centric Approach. *IEEE Network*. 2024;38(5):29-36. [IEEE](#)
- [8] Fang H, Zhang D, Tan C, Yu P, Wang Y, Li W. Large Language Model Enhanced Autonomous Agents for Proactive Fault-Tolerant Edge Networks. *IEEE INFOCOM Workshops*. 2024:1-2. [IEEE](#)
- [9] Alahmari N, Mehmood R, Alzahrani A, Yigitcanlar T, Corchado JM. Autonomous and Sustainable Service Economies: Data-Driven Optimization of Design and Operations through Discovery of Multi-Perspective Parameters. *Sustainability*. 2023;15:16003. [MDPI](#)
- [10] Zahid AF. The Role of Digital Transformation Strategies in Raising the Operational and Marketing Efficiency of Telecommunications Services in the Kingdom of Saudi Arabia: An Applied Study on the Saudi Telecom Company. *The Arab Journal of Administration*. 2025;45(2):239-256. [The Arab Journal of Administration](#)
- [11] Neffati M, Jbir R. The impact of digitalization and economic diversification on economic growth: evidence from Saudi Arabia. *Technological and Economic Development of Economy*. 2024;30(5):1510-1532. [Technological and Economic Development of Economy](#)
- [12] Banerjee A, Mwanje SS, Carle G. Trust and Performance in Future AI-Enabled, Open, Multi-Vendor Network Management Automation. *IEEE Transactions on Network and Service Management*. 2023;20(2):995-1007. [IEEE](#)
- [13] Batool A, Zowghi D, Bano M. AI governance: a systematic literature review. *AI and Ethics*. 2025;5:3265-3279. [Springer](#)
- [14] Tsamados A, Floridi L, Taddeo M. Human control of AI systems: from supervision to

- teaming. *AI and Ethics*. 2025;5:1535-1548. [Springer](#)
- [15] Cohen MK, Kolt N, Bengio Y, Hadfield GK, Russell S. Regulating advanced artificial agents. *Science*. 2024;384(6691):36-38. [Science](#)
- [16] Papagiannidis E, Mikalef P, Conboy K. Responsible artificial intelligence governance: A review and research framework. *Journal of Strategic Information Systems*. 2025;34(2):101885. [ScienceDirect](#)
- [17] Spring M, Faulconbridge J, Sarwar A. How information technology automates and augments processes: Insights from Artificial-Intelligence-based systems in professional service operations. *Journal of Operations Management*. 2022;68(6-7):592-618. [Wiley](#)
- [18] Yu Q, Zhao N, Li M, Li Z, Wang H, Zhang W, Sui K, Pei D. A survey on intelligent management of alerts and incidents in IT services. *Journal of Network and Computer Applications*. 2024;224:103842. [ScienceDirect](#)
- [19] Green B. The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review*. 2022;45:105681. [ScienceDirect](#)
- [20] Trantzas K, Brodimas D, Agko B, Tziavas GC, Tranoris C, Denazis S, et al. Intent-driven network automation through sustainable multimodal generative AI. *EURASIP Journal on Wireless Communications and Networking*. 2025;2025:42. [Springer](#)
- [21] Natarajan UM, Diddigi RB, Bapat J. RAG-inspired Intent-Based Solution for Intelligent Autonomous Networks. *Proceedings of COMSNETS 2025*. 2025:413-421. [IEEE](#)
- [22] Ficzer D, Hollosi G, Varga P. Beyond Intent Translation: Research Gaps in the Application of Generative AI for Intent-Based Networking. *Proceedings of NOMS 2025*. 2025. [IEEE](#)
- [23] Fawzy AH, Wassif K, Moussa H. Framework for automatic detection of anomalies in DevOps. *Journal of King Saud University - Computer and Information Sciences*. 2023;35:8-19. [ScienceDirect](#)
- [24] Poti SP, Stanton CJ. Enabling affordances for AI Governance. *Journal of Responsible Technology*. 2024;18:100086. [ScienceDirect](#)
- [25] Pathak A, Bansal V. AI as decision aid or delegated agent: The effects of trust dimensions on the adoption of AI digital agents. *Computers in Human Behavior: Artificial Humans*. 2024;2(2):100094. [ScienceDirect](#)
- [26] Afroogh S, Akbari A, Malone E, Kargar M, Alambeigi H, et al. Trust in AI: progress, challenges, and future directions. *Humanities and Social Sciences Communications*. 2024;11:1568. [Nature](#)
- [27] Kim A, Sachdeva A, Dennis AR. From self-service to AI-assisted service: A mixed-method study of IT support service provision using search tools and chatbots. *International Journal of Information Management*. 2025;84:102938. [ScienceDirect](#)
- [28] Faulconbridge J, Sarwar A, Spring M. How Professionals Adapt to Artificial Intelligence: The Role of Intertwined Boundary Work. *Journal of Management Studies*. 2025;62(5):1991-2024. [Wiley](#)
- [29] Yu H, Rahimi H, Janz C, Wang D, Yang C, Zhao Y. A Comprehensive Framework for Intent-Based Networking, Standards-Based and Open-Source. *Proceedings of NOMS 2023*. 2023. [IEEE](#)
- [30] Zhang C, Hu D, Yang T. Research of artificial intelligence operations for wind turbines considering anomaly detection, root cause analysis, and incremental training. *Reliability Engineering & System Safety*. 2024;241:109634. [ScienceDirect](#)