



FROM DATA PIPELINES TO DECISION PIPELINES: ENGINEERING THE INFRASTRUCTURE FOR AGENTIC AI

Computer Science

Aditya Ranjan

FNU Independent Researcher

ABSTRACT

The rapid emergence of agentic Artificial Intelligence (AI) systems, in which autonomous software agents plan, reason, invoke tools, and act with limited human intervention, has shifted the central engineering challenge in enterprise AI from moving and transforming data to governing autonomous decisions. Traditional data pipelines were built to extract, transform, and load information for downstream analytics or model training. Agentic AI systems require a further layer, described in this study as the decision pipeline, that manages context assembly, orchestration of multiple specialized agents, tool invocation, memory, guardrails, observability, and human-in-the-loop escalation for actions taken autonomously in production environments. This study examines the engineering infrastructure required to support agentic AI, evaluating its impact on decision latency, orchestration reliability, observability, and organizational trust in autonomous systems. The research also explores AI engineers' and data professionals' perceptions of infrastructure readiness for agentic deployment, along with the opportunities and challenges associated with operationalizing decision pipelines. A descriptive research design was adopted using both primary and secondary data. Primary data were collected through a structured questionnaire administered to AI/ML engineers, data platform architects, and MLOps practitioners, while secondary data were drawn from peer-reviewed literature, arXiv preprints, industry technical reports, and credible online sources on agentic AI, orchestration frameworks, and MLOps infrastructure. Collected data were analyzed using percentage analysis, mean analysis, and graphical representation to assess the maturity of agentic infrastructure practices. The findings indicate that organizations moving from batch data pipelines to real-time decision pipelines experience measurable gains in automation coverage and decision speed, but continue to face challenges around orchestration complexity, authority boundaries between agents and humans, observability of multi-agent reasoning, and governance of autonomous actions. The study concludes that engineering reliable decision pipelines requires treating orchestration, memory, guardrails, and observability as first-class infrastructure layers rather than incidental add-ons to existing data platforms, enabling organizations to deploy agentic AI that is auditable, controllable, and aligned with business risk tolerance.

KEYWORDS

I. INTRODUCTION

Artificial Intelligence has moved through several distinct infrastructure eras. The first era centered on data pipelines: extract-transform-load (ETL) systems, data warehouses, and feature stores built to move information reliably from source systems to analytics dashboards and, later, to model training jobs. The second era, commonly described as MLOps, added versioning, continuous integration and deployment, and monitoring around individual predictive models so that a trained model could be shipped, served, and retrained safely in production [4], [11]. A third era is now emerging around agentic AI, in which large language model (LLM)-based agents plan multi-step tasks, call external tools and application programming interfaces, retrieve and update memory, and take actions with only partial human supervision [7]. This shift changes the fundamental unit of infrastructure concern from a data record or a prediction to a decision: a bounded unit of autonomous reasoning and action that must be logged, explained, and, where necessary, reversed.

Traditional data pipelines were engineered around throughput, schema consistency, and batch or streaming latency. Decision pipelines inherit those requirements but add new obligations: context engineering to assemble the right information for an agent at the right time, orchestration logic to coordinate multiple specialized agents and language models, tool-use frameworks that let agents safely call external systems, memory layers that persist state across a session or across tasks, and guardrail and escalation mechanisms that route uncertain or high-stakes decisions to a human [7], [10]. Industry analyses of production-grade agentic workflows note that organizations increasingly separate a data-plane layer, where agents perform localized interventions such as retrieval or patch generation, from a control-plane layer, where changes to policy, deployment configuration, and approval gates are made, with human oversight concentrated at the control plane [5].

The growing use of agentic systems for operational tasks such as customer support resolution, claims processing, credit risk assessment, and CI/CD automation has made decision-pipeline engineering a strategic priority rather than an experimental exercise. Reports on enterprise agentic deployments describe agents resolving a majority of routine cases autonomously while routing novel or high-risk cases to human operators, which depends on infrastructure that can classify decision risk in real time and act on that classification [7]. Similarly, studies of agentic credit-risk systems describe multi-agent architectures combining reinforcement learning, natural-language reasoning, explainability modules, and continuous feedback loops, while cautioning that model drift, interpretability of high-dimensional

data, and regulatory uncertainty remain open infrastructure challenges [4].

Despite these advantages, engineering decision pipelines introduces challenges that data pipelines rarely faced: non-determinism in agent reasoning, the difficulty of tracing why a multi-agent system reached a given action, the risk of cascading errors across chained tool calls, and the governance question of how much operational authority should be delegated to an autonomous system versus retained by a human [5], [7]. Researchers emphasize that agentic infrastructure should be designed for bounded autonomy, in which agents operate within explicit constraints and recourse mechanisms rather than being granted unrestricted control-plane authority [5].

Digital transformation initiatives across data engineering and platform teams are accelerating the convergence of data infrastructure and AI infrastructure. Industry commentary describes 2025-2026 as the period in which static ETL pipelines give way to context-aware, self-monitoring data systems that feed agentic applications directly, with data quality, observability, and orchestration treated as a single continuum rather than separate concerns [12]. Organizations that engineer this continuum deliberately are better positioned to deploy agentic AI safely and to capture productivity gains, while those that layer agents onto legacy pipelines without new observability and governance tooling report higher rates of production failure and rework [11], [13].

Therefore, this study aims to examine the engineering infrastructure required to support agentic AI by evaluating its impact on decision latency, orchestration reliability, observability, and organizational trust, while identifying the opportunities and challenges associated with the transition from data pipelines to decision pipelines and offering recommendations for building agentic infrastructure that is efficient, auditable, and appropriately governed.

Research Objectives

The primary objective of this study is to examine the engineering infrastructure that underlies agentic AI systems and to evaluate how the transition from conventional data pipelines to decision pipelines affects automation coverage, decision speed, and reliability. The study aims to assess how infrastructure components such as orchestration frameworks, context and memory layers, tool-use gateways, and observability platforms influence agent performance in production. It also seeks to identify the key benefits and challenges associated with agentic infrastructure adoption, including orchestration complexity,

authority boundaries, explainability, and governance of autonomous decisions. Furthermore, the research aims to evaluate the influence of decision-pipeline maturity on organizational trust and operational risk while providing practical recommendations for engineering agentic infrastructure that balances automation with human oversight.

Research Methodology

This study adopts a descriptive research design to examine the engineering infrastructure required for agentic AI. The research is based on both primary and secondary data. Primary data are collected through a structured questionnaire administered to AI/ML engineers, data platform architects, and MLOps practitioners involved in designing, deploying, or operating agentic AI systems. The questionnaire focuses on aspects such as awareness of agentic infrastructure concepts, the perceived contribution of decision pipelines to operational efficiency, the effectiveness of orchestration frameworks, the reliability of agent decision pipelines in production, and overall satisfaction with agentic infrastructure adoption. Secondary data are collected from peer-reviewed journal articles, arXiv preprints, technical industry reports, and credible online sources related to agentic AI, data engineering, and MLOps.

A convenience sampling technique is employed to select the respondents, with a sample size of 112 respondents drawn from engineering and data-platform teams across technology, financial services, and e-commerce organizations. The collected data are organized, classified, and analyzed using statistical tools such as percentage analysis, mean analysis, and graphical representations including bar charts. The analysis helps evaluate the maturity of agentic infrastructure practices and identifies the factors influencing adoption, reliability, and organizational trust in autonomous decision-making. Based on the findings, recommendations are provided to strengthen orchestration design, observability, and governance in agentic AI infrastructure.

II. Review of Literature

Wang, Y., & Zhang, L. (2026):

This study introduces an LLM-powered comparative pipeline for analyzing the governance of agent interoperability protocols, applying it to a permissionless on-chain standard and a corporate-led standard. Using automated annotation, topic modelling, and multi-layer network analysis across thousands of governance participation records, the authors find that governance form shapes thematic priorities even though both regimes show comparable participation inequality, offering infrastructure-level insight into how agentic standards are actually negotiated and controlled in practice [1].

Anonymous Authors, arXiv 2512.23737 (2026):

This paper examines policy-aware agentic control of cloud data pipelines, showing that agentic systems most improve reliability and efficiency when deployed as governed control layers over existing pipeline orchestration rather than as unconstrained autonomous replacements. The study highlights that distributed, compliance-sensitive environments benefit from agents that operate within explicit policy boundaries, reinforcing the case for a distinct governance layer between data infrastructure and autonomous decision-making [2].

Anonymous Authors, arXiv 2601.00818 (2026):

This research proposes a multi-agent system for autonomous, explainable, real-time credit-risk decision-making, combining reinforcement learning, natural-language reasoning, explainability modules, and continuous feedback pipelines. The findings show improved decision speed, transparency, and responsiveness relative to traditional scoring models, while identifying model drift, interpretability of high-dimensional data, regulatory uncertainty, and infrastructure limitations in low-resource settings as continuing challenges for decision-pipeline engineering [3].

Anonymous Authors, arXiv 2605.07062 (2026):

This paper reframes agentic CI/CD as a problem of authority transfer rather than task performance, distinguishing data-plane authority, such as localized patch generation and test reruns, from control-plane authority, such as changes to deployment policy and approval gates. Drawing on research prototypes and industrial platforms, the authors show that current systems operate mainly at the data plane under bounded autonomy, with safety achieved through surrounding governance infrastructure rather than intrinsic guarantees within the agent itself [4].

Anonymous Authors, arXiv 2512.08769 (2025):

This practical guide outlines a structured engineering lifecycle for designing, developing, and deploying production-grade agentic AI workflows, arguing that agentic systems integrate multiple specialized agents, tool-augmented capabilities, and orchestration logic into dynamic pipelines capable of autonomous decision-making. The authors emphasize that reliability, observability, maintainability, and alignment with safety and governance requirements must be engineered explicitly rather than assumed as by-products of capable underlying models [5].

Anonymous Authors, arXiv 2605.18764 (2026):

This study proposes a controlled, human-in-the-loop agentic framework that guides domain experts through problem definition, compute specification, pipeline generation, and code generation to produce executable AI solutions. Empirical studies show that the framework adapts to diverse problem contexts while producing structured, reusable artefacts, illustrating how staged human oversight can be embedded directly into pipeline-generation infrastructure rather than added afterwards [6].

Mirantis Research Note (2026):

This industry analysis identifies context-length and reasoning bottlenecks, the benefits of specialization through modular agents, and advances in communication protocols as forces driving the shift toward collaborative, multi-agent ecosystems. It reports that governed agentic deployments in customer service settings resolve a large share of routine incidents without human intervention while routing novel or high-risk cases to human operators, and that agentic systems running on properly managed infrastructure improve GPU utilization by routing inference on real-time cost and latency signals [7].

Business and Technology Press on MLOps Maturity (2025-2026):

Multiple industry reports describe the convergence of data engineering and AI engineering into a single operational continuum, noting that a majority of production machine-learning teams rely on lifecycle-management platforms for experiment tracking and deployment, that compliance frameworks are pushing organizations toward stronger monitoring and explainability, and that a systematic literature review of MLOps practice identifies unified versioning of code, data, and model artefacts as among the most consistently validated engineering practices for production AI systems [8], [9].

Ehtesham, A., Singh, A., & Kumar, S. (2025):

This survey compares the leading agent-interoperability protocols, including the Model Context Protocol, Agent Communication Protocol, Agent-to-Agent Protocol, and Agent Network Protocol, describing a shift toward lightweight, standardized protocols that enable dynamic discovery, secure communication, and decentralized collaboration across heterogeneous agent systems. The review positions protocol-level standardization as a foundational infrastructure layer beneath any decision pipeline that spans more than one agent or tool provider [16].

Anonymous Authors, arXiv 2601.13671 (2026):

This paper consolidates the technical composition of orchestrated multi-agent systems into a unified architectural framework integrating planning, policy enforcement, state management, and quality operations into a coherent orchestration layer. It provides a detailed treatment of how the Model Context Protocol standardizes agent access to tools and context while a peer-to-peer agent protocol governs coordination, negotiation, and delegation, together forming an auditable, policy-compliant communication substrate for enterprise adoption [17].

Anonymous Authors, arXiv 2606.30546 (2026):

This work presents a specification-driven validation framework for reliable multi-agent systems, drawing on open telemetry standards and agent-observability platforms to argue that reliability in agentic infrastructure depends on structured tracing of every model call, tool invocation, and coordination step rather than on evaluating final outputs alone. The framework illustrates how observability tooling built for conventional software must be extended to capture agent-specific failure modes [18].

III. Data Analysis & Interpretation

1. Awareness of Agentic AI Infrastructure Concepts

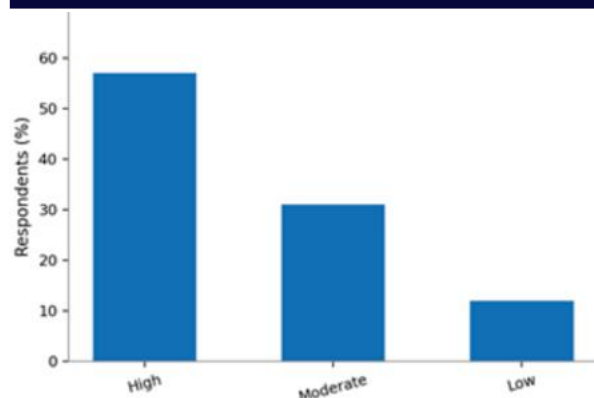


Fig 1: Awareness of Agentic AI Infrastructure Concepts

Interpretation: A majority of respondents report high awareness of agentic AI infrastructure concepts such as orchestration frameworks, tool-use gateways, and decision pipelines, reflecting the rapid mainstreaming of agentic engineering practice among data and ML professionals.

2. Decision Pipelines Improve Operational Efficiency

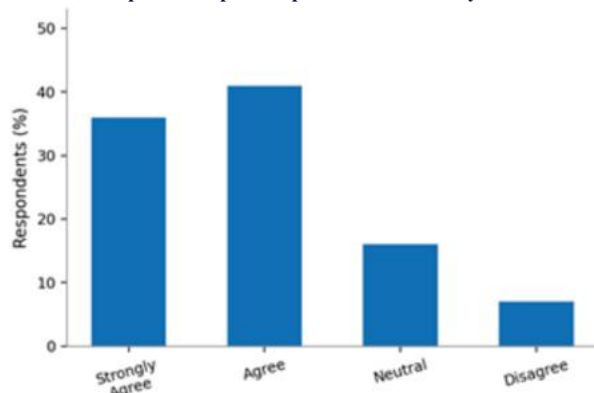


Fig 2: Decision Pipelines Improve Operational Efficiency

Interpretation: Most respondents agree that structured decision pipelines, built on top of existing data pipelines, meaningfully reduce manual intervention and shorten the time taken to complete multi-step operational tasks.

3. Effectiveness of Orchestration Frameworks

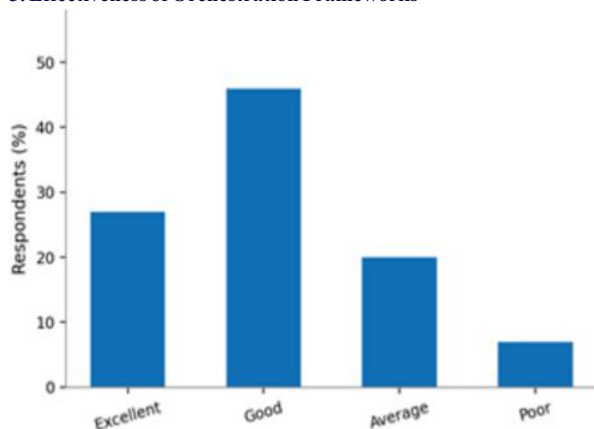


Fig 3: Effectiveness of Orchestration Frameworks in Managing Agent Pipelines

Interpretation: Orchestration frameworks that coordinate multiple agents, tools, and memory stores are rated as effective by most respondents, though a notable share still reports only average performance, pointing to room for improvement in coordination logic.

4. Reliability of Agent Decision Pipelines in Production

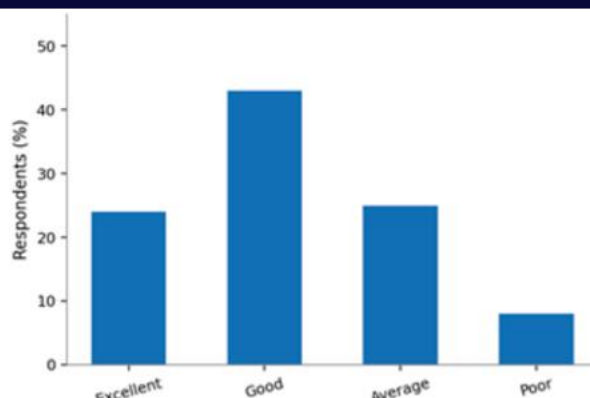


Fig 4: Reliability of Agent Decision Pipelines in Production

Interpretation: Reliability ratings are positive overall, but a larger proportion of respondent's rate production reliability as average or poor compared with orchestration effectiveness alone, suggesting that failure modes emerge more often at the level of end-to-end decision execution than within individual orchestration components.

5. Overall Satisfaction with Agentic AI Infrastructure Adoption

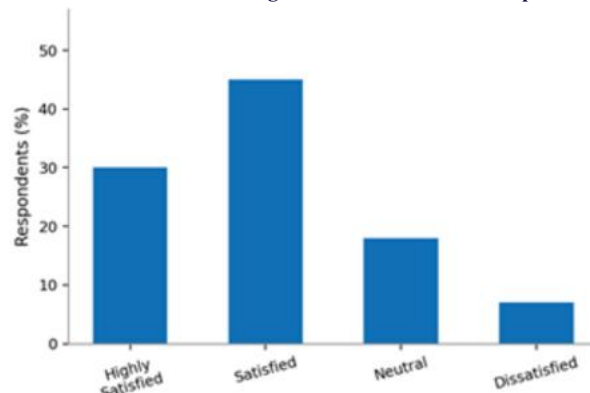


Fig 5: Overall Satisfaction with Agentic AI Infrastructure Adoption

Interpretation: Respondents expressed overall satisfaction with agentic AI infrastructure adoption, though the share reporting neutral or dissatisfied responses is higher than for data-pipeline-only initiatives, consistent with the added complexity of governing autonomous decisions.

IV. Findings

The findings of the study reveal that engineering decision pipelines on top of existing data pipelines meaningfully improves the speed and consistency of operational decision-making. Most respondents agreed that orchestration frameworks, context and memory layers, and tool-use gateways reduce manual coordination effort and allow multi-step tasks to be completed with less direct human involvement. The study also found that observability tooling, including tracing of agent reasoning steps and tool calls, is a strong predictor of respondents' trust in autonomous decisions, with teams that invest in structured logging and tracing reporting higher confidence in production reliability. Additionally, organizations that define explicit boundaries between agent authority and human approval, particularly for actions affecting external systems or financial outcomes, report fewer incidents of unintended or cascading agent actions.

The study further identified certain challenges associated with the transition from data pipelines to decision pipelines, including the difficulty of debugging non-deterministic multi-agent reasoning, inconsistent standards for context engineering across teams, the operational cost of maintaining low-latency memory and retrieval layers, and continuing uncertainty about regulatory expectations for autonomous decision-making in regulated industries. Respondents emphasized that while agentic infrastructure improves efficiency, high-stakes or irreversible actions should continue to route through human review, echoing the bounded-autonomy design pattern found in current literature. Overall, the findings indicate that decision pipelines are most effective when treated as governed infrastructure, layered

deliberately on top of well-engineered data pipelines, rather than as an unconstrained extension of existing automation.

V. CONCLUSION

The study concludes that engineering the infrastructure for agentic AI represents a natural but non-trivial evolution beyond conventional data pipelines and MLOps practice. Decision pipelines, comprising orchestration, context and memory management, tool-use gateways, observability, and guardrail mechanisms, enable organizations to automate multi-step operational decisions while preserving the ability to audit, explain, and, where necessary, override autonomous actions. The findings indicate that integrating decision-pipeline infrastructure with existing data platforms improves decision speed and reduces manual workload, while requiring new engineering discipline around non-determinism, authority boundaries, and end-to-end observability that traditional data infrastructure did not need to address.

However, the study also concludes that agentic infrastructure should complement rather than replace human oversight in operational decision-making. Challenges such as orchestration complexity, opaque multi-agent reasoning, and unresolved governance and regulatory questions require careful, deliberate infrastructure design and continuous monitoring. Organizations should treat decision pipelines as first-class infrastructure investments, with explicit control-plane governance, rather than as incidental extensions of existing ETL or MLOps tooling. Overall, the research highlights that a balanced approach, combining well-engineered decision pipelines with human-in-the-loop governance, can help organizations deploy agentic AI that is efficient, trustworthy, and appropriately accountable for the decisions it makes.

VI. Future Scope

The scope for future research on engineering the infrastructure for agentic AI is extensive as agent capabilities and organizational adoption continue to mature. Future studies can examine the long-term impact of decision-pipeline infrastructure on incident rates, regulatory compliance, and workforce roles across industries with differing risk tolerance, such as healthcare, finance, and manufacturing. Comparative research across organizations with varying orchestration maturity could clarify which infrastructure investments most reduce production failures in agentic systems. Researchers may also explore emerging standards for agent interoperability and communication protocols, explainable multi-agent reasoning, standardized context-engineering practices, and formal methods for verifying bounded-autonomy guarantees. In addition, future studies can investigate the evolving regulatory landscape for autonomous decision-making, cost-performance trade-offs in memory and retrieval infrastructure, and longitudinal studies with larger, cross-industry samples to build a more complete picture of how decision pipelines should be engineered, governed, and evaluated as agentic AI moves further into core business operations.

VII. REFERENCES

- [1] Wang, Y., & Zhang, L. (2026). Agentic analysis for agentic infrastructure: An LLM-powered pipeline for comparative governance of DAO and corporate AI protocols. arXiv:2606.26203. <https://arxiv.org/abs/2606.26203>
- [2] Governing cloud data pipelines with agentic AI (2026). arXiv:2512.23737. <https://arxiv.org/pdf/2512.23737>
- [3] Agentic AI for autonomous, explainable, and real-time credit risk decision-making (2026). arXiv:2601.00818. <https://arxiv.org/pdf/2601.00818>
- [4] From assistance to agency: Rethinking autonomy and control in CI/CD pipelines (2026). arXiv:2605.07062. <https://arxiv.org/pdf/2605.07062>
- [5] A practical guide for designing, developing, and deploying production-grade agentic AI workflows (2025). arXiv:2512.08769. <https://arxiv.org/html/2512.08769>
- [6] From intent to AI pipelines: A controlled agentic framework for non-AI expert scientists (2026). arXiv:2605.18764. <https://arxiv.org/html/2605.18764v1>
- [7] Understanding agentic AI infrastructure (2026). Mirantis Technical Blog. <https://www.mirantis.com/blog/agentic-ai-infrastructure/>
- [8] MLOps in 2026: From MLflow to LLMops — the complete guide to shipping AI in production (2026). Medium/CodeX. <https://medium.com/codex/mlops-in-2026-from-mlflow-to-llmops-the-complete-guide-to-shipping-ai-in-production-0024955b70c4>
- [9] What is MLOps? 2026 enterprise guide to scaling production AI (2026). Modern Data 101 Blog. <https://www.moderndata101.com/blogs/best-practices-for-mlops-for-scaling-ai>
- [10] AI tech stack 2026: Frameworks, MLOps & IDEs guide (2026). Kelton Technical Blog. <https://www.kelton.com/kelton-tech-blog/ai-tech-stack-2026>
- [11] Enterprise machine learning platforms: A buyer's guide for 2026 (2026). Dataiku. <https://www.dataiku.com/stories/blog/enterprise-machine-learning-platforms-how-to-choose>
- [12] AI-powered data orchestration: How intelligent pipelines are redefining data engineering in 2026 (2026). BOT Consulting. <https://www.botconsulting.io/post/ai-powered-data-orchestration-how-intelligent-pipelines-are-redefining-data-engineering-in-2026>
- [13] MLOps services: The 2026 strategy for scaling enterprise AI (2026). Moog Labs. <https://www.mooglabs.com/blog/mlops-enterprise-strategies>
- [14] DevOps and MLOps pipelines guide 2026 (2026). GitNexa Blog. <https://www.gitnexa.com/blogs/devops-for-ai-ml-projects>
- [15] MPAC: A multi-principal agent coordination protocol for interoperable multi-agent collaboration (2026). arXiv:2604.09744. <https://arxiv.org/pdf/2604.09744>
- [16] Ehtesham, A., Singh, A., & Kumar, S. (2025). A survey of agent interoperability protocols: Model Context Protocol (MCP), Agent Communication Protocol (ACP), Agent-to-Agent Protocol (A2A), and Agent Network Protocol (ANP). arXiv:2505.02279. <https://arxiv.org/html/2505.02279v1>
- [17] The orchestration of multi-agent systems: Architectures, protocols, and enterprise adoption (2026). arXiv:2601.13671. <https://arxiv.org/abs/2601.13671>
- [18] MAS-Lab: A specification-driven validation framework for reliable multi-agent systems (2026). arXiv:2606.30546. <https://arxiv.org/pdf/2606.30546>
- [19] How to orchestrate multi-agent AI systems at scale in 2026 (2026). Atlan. <https://atlan.com/know/multi-agent-system-orchestration/>
- [20] Best multi-agent orchestration frameworks in 2026 (2026). TrueFoundry. <https://www.truefoundry.com/blog/multi-agent-orchestration-frameworks>