Journal of Economics, Finance and Accounting Studies (JEFAS)
ISSN:
DOI:
Journal Homepage: www.al-kindipublisher.com/index.php/jefas

JEFAS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

A Deterministic Trajectory-Level Evaluation Framework for Learning-Based Agentic Systems

Prasad Maderamitla[1] and Subba Rao Katragadda[2]
[12]*Independent researcher, California, USA*
Corresponding Author: Prasad Maderamitla, **E-mail**: prasad.madera@gmail.com

| ABSTRACT

In the recent past, learning-based agentic systems are being increasingly used to tackle complex decision environments where reliability, transparency, and compliance with governance are key. However, the dominant evaluation strategies are mostly outcome-centric, with a focus on aggregate performance measures like accuracy, rewards, or task completion rates. These evaluation strategies provide little information on the internal decision-making processes that lead to the observable outcomes, especially for reasoning-based agentic systems with multiple steps. This paper proposes a Deterministic Trajectory Level Evaluation Framework (DTLEF) for learning-based agentic systems. The proposed framework transforms the evaluation paradigm from traditional outcome-oriented metrics to an evaluation process that focuses on the trajectories of states and actions within controlled execution conditions. The DTLEF integrates standardized states, comprehensive action trace logging, deterministic replay validation, and behavior verification against governance constraints. The evaluation process, in which the agentic system runs in controlled inference mode and the trajectory traces are compared, enables the identification of policy instabilities, reasoning drift, and non-deterministic behavior. Moreover, the evaluation process also ensures that the trajectories are consistent with the predefined constraints. Unlike traditional performance metrics, the proposed framework does not depend upon empirical data or domain-specific metrics. Instead, it provides a methodology for evaluating agentic systems, including autonomous decision pipelines, tool-enhanced language agents, and cyber-physical control systems, at an architecture level. The proposed framework increases transparency, reproducibility, and compliance without modifying training procedures. This research provides a scalable and domain-independent evaluation methodology for validating learning-based autonomous agents in environments where behavioral reliability is as important as functional performance by formally defining trajectory-level determinism as a primary evaluation criterion.

| KEYWORDS

Agentic Systems, Trajectory-Level Evaluation, Deterministic Replay, Learning-Based Autonomous Agents, AI Governance and Reliability

| ARTICLE INFORMATION

ACCEPTED: 20 January 2026 **PUBLISHED:** 17 February 2026 **DOI:** 10.32996/jefas.2026.8.3.3

1. Introduction

Learning-based agentic systems, which are autonomous agents designed to perform perception, reasoning, planning, and action, have made tremendous progress in recent years due to the advances in deep learning, reinforcement learning, and large language models. Such agentic systems are increasingly being used in safety-critical and mission-critical applications, ranging from autonomous manufacturing cells to financial decision-making systems, health diagnostics, and cyber-physical systems. While agentic systems have shown tremendous promise in terms of their functional performance, the most important challenge in the field of agentic systems has been the proper evaluation of the agent's behavior [1].

The most commonly used techniques for evaluating agentic systems are based on task-level performance metrics, which include reward optimization, accuracy, precision-recall curves, and completion rates. However, it has been found that the decision-making process of the agent, which ultimately determines the agent's behavior, is often opaque in the case of agentic systems, especially those based on learning-based techniques. Such opaqueness of the decision-making process of the agent has been a major challenge in the field of agentic systems [2].

These developments in XAI and VAI emphasize the importance of evaluation approaches that are effective at the trajectory level as well as the output level, allowing for the inspection and validation of sequences of intermediate states, actions, and reasoning. However, the majority of the existing approaches are stochastic and do not guarantee determinism across multiple executions of the same evaluation process [3].

This paper presents the concept of the Deterministic Trajectory-Level Evaluation Framework (DTLEF), which is specific to the evaluation of learning-based agentic systems. The DTLEF shifts the focus of evaluation approaches from outcome-based measures to trajectory consistency, state-action determinism, rule satisfaction, and reproducibility in well-defined execution environments, with the help of the adoption of trajectory decomposition and deterministic replay validation.

The primary contributions of this work include:

- A formal definition of trajectory-level evaluation for learning-based agents.
- A deterministic replay architecture for verifying behavioral consistency.
- A structured evaluation metric set focusing on state alignment, policy determinism, and rule conformance.
- A theoretical validation framework applicable across autonomous decision systems.

2. Literature review

2.1 Evaluation of Learning-Based Agents

The evaluation paradigms for learning-based agents have, in most cases, been based on the reinforcement learning paradigm, where the cumulative expected rewards are the main performance metrics. Some of the common and widely used performance metrics, such as episodic return, regret minimization, and success rate, have been the mainstay of most of the literature. These metrics, although providing a measure of performance, do not offer much insight into the underlying decision processes [4][5].

The recent trend in the study of learning-based agents has been to emphasize the aspect of explainability. Saliency mapping, attention visualization, and post-hoc explanation models have been some of the techniques used to explain the underlying decision processes. These techniques, however, have been observational in nature and do not offer a deterministic guarantee [6].

2.2 Agentic Systems and Multi-Step Reasoning

Agentic systems are differentiated from other models based on sequential reasoning, retention of memory, interaction with the environment, and adaptive planning. The use of large language models with tool invocation abilities introduces multi-step reasoning chains. These systems are defined in dynamic state spaces, making it complicated to evaluate the trajectories [7].

Currently, the verification methodologies used for autonomous systems are mainly based on formal methods, model checking, and runtime monitoring. Although these methodologies are effective for rule-based and deterministic systems, they are faced with the challenge of scalability when used with learning-based systems because of the stochastic nature of the models and the nonlinear mapping of the models [8].

2.3 Determinism and Reproducibility in AI

Determinism in artificial intelligence systems has also received attention for safety-critical applications. Deterministic replay is used for debugging distributed systems and robotics simulation environments. In the context of reinforcement learning, issues with reproducibility arise due to sensitivity to random seeds, environment randomness, and policy exploration.

However, most deterministic works focus on training reproducibility rather than the evaluation of the behavior of the agent. There is limited theoretical work done on deterministic evaluation systems that are independent of training [9].

2.4 Research Gap

The literature has identified three major limitations:

- Excessive reliance on aggregate performance measures.

- Lack of adequate introspection mechanisms for trajectory levels.
- Lack of deterministic replay validation for agentic behavior.

This work aims to address all these shortcomings by proposing a deterministic evaluation architecture.

3. Methodology

The proposed Deterministic Trajectory-Level Evaluation Framework (DTLEF) for the evaluation of learning-based agentic systems considers the complete decision-making process rather than the final task outcomes. The proposed evaluation methodology is based on structured state representation, action recording, deterministic replay validation, and governance-oriented behavioral analysis. The proposed DTLEF is an evaluation layer that is independent of the training process.

3.1 Structured State Representation

The first part of the framework is the formalization of the representation of the system states. In the context of agentic systems, states are generally made up of environmental inputs, internal memory, intermediate reasoning context, and task stage metadata. In the interest of promoting reproducibility and conceptual clarity, the DTLEF proposes the standard schema for states that includes all information relevant to decision-making at every step.

This representation eliminates the ambiguity of the reconstruction of the state and ensures that each state transition can be exactly recorded and replayed. By making the intermediate reasoning explicit, the proposed framework enables the inspection of the trajectories rather than the evaluation based solely on the black-box output.

3.2 Comprehensive Action Trace Logging

DTLEF maintains a systematic record of all the high-level actions taken by the agent during the course of the task execution. The variety of actions ranges from tool invocations, planning decisions, memory modifications, and environment interactions. Instead of restricting the recording of actions to terminal outputs, the framework retains the entire sequence of state-action transitions.

Each action entry retains the state of the environment prior to the action and the state of the environment after the action. This provides the evaluators with the ability to analyze the development of decision-making over time. The recording of actions occurs in the passive role of the framework and does not alter the decision-making process of the agent.

3.3 Deterministic Replay Validation

One of the key aspects of the methodology is the use of deterministic replay. In the evaluation process, the agent is subjected to a controlled inference process where non-deterministic sampling is minimized. The inputs and the context are replayed several times to evaluate the consistency of the state and action.

If the results are the same after several iterations, it indicates that the system has achieved trajectory level determinism under controlled conditions. In the event of a deviation, the framework is able to identify the exact step where the deviation occurs.

Deterministic replay ensures that the evaluation process considers the reliability of the process as well as the performance effectiveness.

3.4 Trajectory Consistency and Governance Alignment

Besides the issue of replay stability, DTLEF also examines the integrity of the trajectory and the satisfaction of the rules. It compares the repeated execution of the trajectories in order to detect policy drift, unstable reasoning paths, and sensitivity to small variations in the context.

Additionally, the state-action transitions are checked against the defined operational constraints, be it safety regulations, organizational policies, or boundaries of the domain. It should be noted that determinism, by itself, is not enough unless it is coupled with system-wide constraints.

3.5 Evaluation Protocol

The methodology is performed within a standardized evaluation protocol, which defines controlled inference mode, environmental conditions, and contextual inputs. The framework is a theoretical one, architecture-centric, and domain-independent, as it does not rely on empirical data sets.

DTLEF, through its combination of structured trajectory logging, deterministic replay, and governance-aware validation, provides a systematic and reproducible approach for evaluating learning-based agentic systems at a process level rather than an outcome level[10].

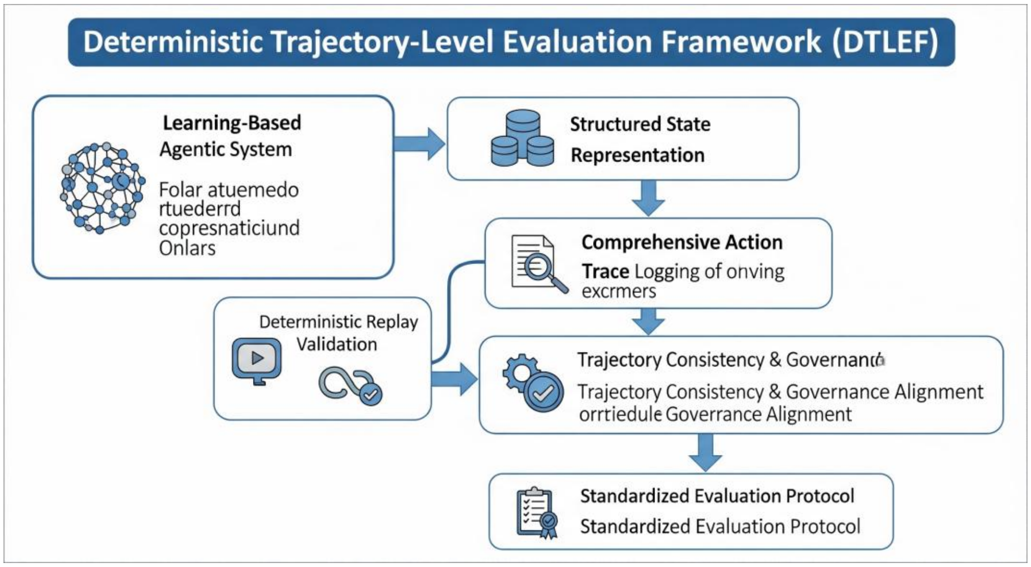

Fig 1. DTLEF architecture Framework

The figure illustrates the conceptual architecture of the Deterministic Trajectory-Level Evaluation Framework (DTLEF) as an external evaluative framework in relation to a learning-based agentic system. The diagram illustrates a workflow that is process-oriented and sequential in nature, starting from the Learning-Based Agentic System and moving through the structured evaluation components in a linear fashion. First, the conditions of the system and the environmental conditions are represented using the Structured State Representation. Moving ahead in the workflow, the Comprehensive Action Trace Logging component logs all the actions taken throughout the execution process. The Deterministic Replay Validation component re-executes the agentic system in a deterministic manner to check the consistency of the trajectories. The Trajectory Consistency & Governance Alignment component aligns the multiple execution traces of the system and checks the consistency of the trajectories. Finally, the workflow ends at the Standardized Evaluation Protocol component, which emphasizes the importance of the framework in terms of ensuring the reliability and transparency of the system.

4. Discussion

The DTLEF framework shifts the evaluation paradigm from probabilistic outcome-based assessment to deterministic process-level validation. This has several implications:

4.1 Transparency and Explainability

By logging structured trajectories, decision pathways become inspectable. Instead of explaining isolated outputs, stakeholders can trace sequential reasoning steps.

4.2 Safety and Compliance

In regulated domains such as industrial automation or finance, deterministic replay ensures that identical inputs produce identical decisions. This reduces liability and improves auditability.

4.3 Robustness Enhancement

Trajectory divergence metrics enable early detection of unstable policies. Systems with high TDI values may exhibit sensitivity to minor perturbations, indicating the need for policy refinement.

4.4 Limitations

The framework assumes controlled evaluation environments. In real-world stochastic environments, perfect determinism may not always be achievable. Additionally, deterministic inference configurations may not reflect natural exploration behavior.

4.5 Scalability Considerations

Trajectory storage may become computationally intensive for long-horizon agents. Compression and hierarchical abstraction techniques may be required for large-scale systems.

5. Conclusion

The present article introduces a Deterministic Trajectory-Level Evaluation Framework (DTLEF) for the evaluation of learning-based agentic systems. In contrast to typical evaluation paradigms, where the overall performance of the system is considered, the DTLEF introduced here emphasizes the consistency of trajectories, deterministic validation of replays, and rule-constrained behavioral evaluation.

The formalization of trajectory decomposition and the introduction of structured metrics, such as State Transition Stability and Trajectory Divergence Index, improve the overall reproducibility, transparency, and governance of the system. Even if the DTLEF is, to a certain extent, a purely theoretical approach, it provides a highly scalable and domain-independent architecture for the evaluation of autonomous decision systems.

Possible future developments could involve the extension of the DTLEF to a hybrid deterministic-probabilistic evaluation model and its application to collaborative systems and cyber-physical systems.

References

[1]. Bengio Y (2025) The First International AI Safety Report. SuperIntelligence - Robotics - Safety & Alignment 2:. https://doi.org/10.70777/si.v2i2.14755

[2]. Milani S, Topin N, Veloso M, Fang F (2024) Explainable Reinforcement Learning: A Survey and Comparative Review. ACM Comput Surv 56:. https://doi.org/10.1145/3616864

[3]. Kim W, Park S, In Y, et al (2025) Beyond the Final Answer: Evaluating the Reasoning Trajectories of Tool-Augmented Agents

[4]. Itu A (2025) Industrial Scheduling in the Digital Era: Challenges, State-of-the-Art Methods, and Deep Learning Perspectives. Applied Sciences (Switzerland) 15

[5]. Glanois C, Weng P, Zimmer M, et al (2024) A survey on interpretable reinforcement learning. Mach Learn 113:. https://doi.org/10.1007/s10994-024-06543-w

[6]. Müller R (2025) How Explainable AI Affects Human Performance: A Systematic Review of the Behavioural Consequences of Saliency Maps. Int J Hum Comput Interact 41:. https://doi.org/10.1080/10447318.2024.2381929

[7]. Sumers TR, Yao S, Narasimhan K, Griffiths TL (2024) Cognitive Architectures for Language Agents. Transactions on Machine Learning Research 2024:

[8]. Langford MA, Chan KH, Fleck JE, et al (2023) MoDALAS: addressing assurance for learning-enabled autonomous systems in the face of uncertainty. Softw Syst Model 22:. https://doi.org/10.1007/s10270-023-01090-9

[9]. Nagarajan P, Warnell G, Stone P (2019) Deterministic Implementations for Reproducibility in Deep Reinforcement Learning

[10]. Prasad Maderamitla, & Subba Rao Katragadda. (2026). Observability for LLM apps: what to log, privacy-safe telemetry, KPIs. Frontiers in Computer Science and Artificial Intelligence, 5(4), 10-14. https://doi.org/10.32996/jcsts.2026.5.4.2