International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Closing the Loop: Real-World Impact of Generative World Model Simulation on AV Safety Validation

Tamoghna Das

Abstract: The transition from open-loop to closed-loop evaluation represents the defining methodological shift in autonomous vehicle safety validation over the past two years. Generative world models have made this transition operationally viable by enabling reactive, controllable, and physically plausible simulation environments at scales that physical test tracks and replay-based simulators cannot match. This article examines four concrete applications of closed-loop world model simulation that have demonstrated measurable safety outcomes: city-scale traffic simulation with full journey evaluation capability (sustaining coherent simulation beyond sixty seconds), novel-trajectory assessment for detecting geometric breakdown at distribution boundaries (with a documented 195.87 percent improvement in novel-trajectory fidelity), reinforcement learning policy hardening in world-model environments (achieving a threefold reduction in collision rate versus imitation-learning baselines), and counterfactual dataset construction for safety-critical training data augmentation (producing a substantial drop in safety violations on challenging scenarios). Drawing on peer-reviewed contributions from 2024 and 2025, the analysis identifies three cross-cutting mechanisms by which closed-loop simulation translates generative fidelity into operational safety improvement: causal coupling that forces behavioral coherence between agent actions and scene evolution, distributional coverage of rare events that fleet logs cannot supply, and reward-aligned policy optimization that targets safety margins rather than perceptual plausibility. The article concludes with a deployment timing framework that conditions the adoption of each application on the maturity of the world model infrastructure already in place, providing a practical decision guide for safety engineering teams evaluating the transition from experimental simulation tooling to production safety pipelines.

Index Terms: *autonomous vehicle safety validation; closed-loop simulation; generative world models; counterfactual data augmentation; reinforcement learning; rare-event coverage; causal coupling*

I. Introduction

Fleet-based autonomous vehicle development has reached a scale at which the fundamental constraint on safety improvement is no longer hardware, compute, or raw driving data. The binding constraint is the capacity to evaluate policy behavior in scenarios that are simultaneously rare, safety-critical, and causally coherent. A collision avoidance maneuver that a policy must execute correctly to prevent a serious crash may occur once in hundreds of millions of driven miles on public roads. No physically realistic test program can produce the volume of such encounters required to train and validate a robust policy.

tamoghnad55@gmail.com

Independent Researcher, USA

ORCID ID: 0009-0004-1308-7050

Generative world models resolve this constraint by synthesizing driving scenarios on demand [7], [8]. The critical distinction between a world model that is useful for safety validation and one that is not is causal coupling: the world model must translate agent actions into scene state updates that are geometrically and behaviorally consistent with those actions [12]. A model that generates visually plausible future frames without incorporating the agent's steering and acceleration outputs into the scene transition produces training data that teaches the policy to ignore the causal structure of the environment, with predictable consequences for real-world performance.

Four applications of closed-loop world model simulation have produced documented safety outcomes. City-scale traffic simulation with full journey evaluation capability addresses the temporal scope limitation of replay-based systems, sustaining

coherent simulation for durations exceeding sixty seconds [10]. Novel-trajectory assessment provides a systematic method for detecting world model failure modes at the boundary of the training distribution, with documented gains of 195.87 percent in novel-trajectory fidelity [4]. Reinforcement learning in world-model environments enables reward-driven policy optimization achieving a threefold reduction in collision rate versus imitation-learning baselines [23]. Counterfactual dataset construction converts near-miss events captured by the fleet into calibrated safety-critical training scenarios, producing a substantial drop in safety violations on challenging benchmarks [11].

The research question this article addresses is: through what mechanisms does closed-loop generative world model simulation translate into measurable autonomous vehicle safety improvement, and what deployment conditions determine whether each application is operationally viable? The answer has both technical and organizational dimensions, and the deployment timing framework developed in Section VIII addresses both.

II. The Validation Gap in AV Safety Testing

Contemporary autonomous vehicle safety validation rests on three overlapping methodologies: structured test track evaluations, replay-based simulation using recorded sensor logs, and statistical analysis of fleet-collected driving data. Each methodology addresses a subset of the safety validation problem, and each has a structural limitation that the others do not compensate for.

Structured test track evaluations provide ground-truth safety outcomes for a controlled set of scenarios, but their scenario coverage is inherently bounded by physical and logistical constraints. The cost of executing a single structured test scenario with human safety drivers, instrumented vehicles, and controlled traffic actors is measured in thousands of dollars. The number of structurally distinct safety-critical scenarios relevant to a mature autonomous driving system is measured in the thousands, making full structured test coverage economically impractical.

Replay-based simulation addresses coverage scale by executing the autonomous system against a large library of recorded sensor logs [1], [9]. The fundamental limitation of replay-based simulation is the absence of agent reactivity: the recorded agents in a logged scenario do not respond to the behavior of the system under test. A policy that avoids a collision by a different margin than the original driver did is evaluated against a ground truth that assumes the original driver's margin, systematically underestimating the collision risk introduced by the policy deviation.

Fleet data analysis provides statistical grounding but operates on a historical record that reflects the behavior of previously deployed system versions, not the system under evaluation. The rare-event scenarios most relevant to safety improvement are those least represented in fleet logs, by definition. These three structural limitations together constitute the validation gap that closed-loop generative world model simulation is designed to close [13], [14].

Table 1. Validation Paradigm Comparison: Structural Capabilities and Limitations

Paradigm	Agent Reactivity	Rare-Event Coverage	Causal Coupling	Scale
Structured test tracks	Full	Low	Full	Very Low
Replay-based simulation	None	Low	None	High
Fleet data analysis	None	Very Low	None	Very High
Open-loop world model	None	Moderate	None	High
Closed-loop world model	Full	High	Full	High

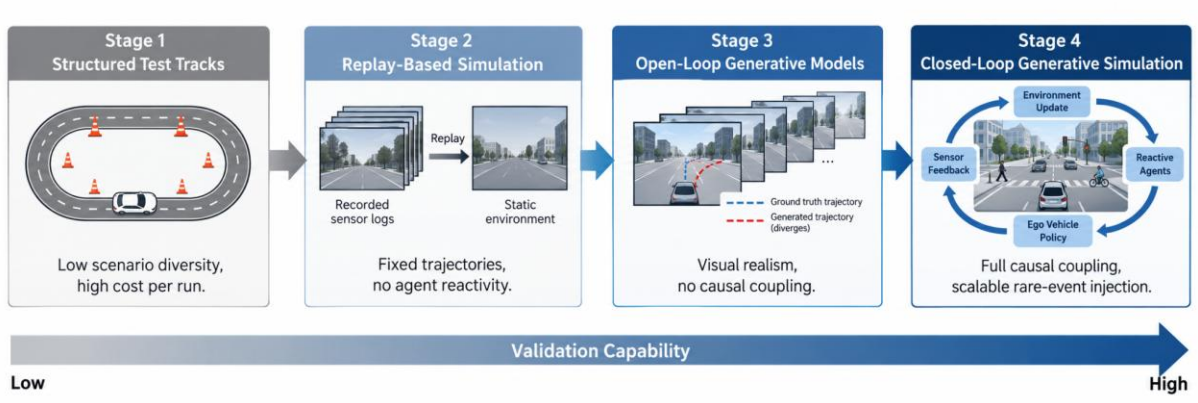


Fig. 1. Evolution of AV safety validation paradigms from structured test tracks to closed-loop generative simulation, illustrating increasing causal fidelity and scenario diversity at each stage.

III. City-Scale Traffic Simulation with Full Journey Evaluation

The temporal scope of world model simulation has been a persistent constraint. Early closed-loop world model deployments could sustain physically plausible simulation for intervals of two to five seconds, sufficient for isolated interaction evaluation but inadequate for assessing policy behavior across the full sequence of decisions that constitutes a driving journey from origin to destination [15], [16].

SceneDiffuser++ [10], presented by Waymo Research at CVPR 2025, addressed this constraint through a joint spatiotemporal diffusion architecture that synthesizes the complete trajectory set for all agents in a scene simultaneously rather than agent-by-agent. The system sustains coherent city-scale simulation for durations exceeding sixty seconds — a substantial advance over prior systems that lost geometric coherence within a few seconds. The joint generation approach enforces behavioral coherence at the scenario level: agent trajectories are consistent with one another by construction, producing traffic configurations that reflect realistic collective behavior rather than superimposed independent agent samples [17]. The CitySim task introduced alongside SceneDiffuser++ defines a full-journey evaluation protocol in which the system under test is evaluated across complete origin-to-destination routes rather than isolated interaction clips, providing a safety assessment that captures the cumulative risk profile of a policy rather than its performance on selected critical moments.

The practical significance of full-journey evaluation is substantial for safety case construction. A policy that performs well on a library of isolated critical interaction scenarios may still exhibit systematic navigation errors, lane discipline failures, or speed profile deviations that accumulate across a full journey into a materially elevated risk profile. City-scale simulation with joint agent generation provides the scenario scope to detect these cumulative failures before fleet deployment [3].

IV. Novel-Trajectory Assessment and Distribution Boundary Detection

World model reliability degrades as the input trajectory diverges from the training distribution. This degradation is not gradual: systems that maintain high geometric fidelity within the training distribution can exhibit abrupt geometric breakdown, including scene element detachment, ground plane instability, and agent bounding box drift, when the ego trajectory exceeds the curvature, speed, or lane position envelope of the training data [6].

ReconDreamer [4], presented at CVPR 2025, quantified this failure mode on the Waymo Open Dataset [9] and proposed online restoration as a mitigation strategy. Rather than completing a novel trajectory in a single forward pass, ReconDreamer iteratively incorporates intermediate synthesized frames into the conditioning context, progressively stabilizing scene geometry as the trajectory diverges from the training distribution. The system reported a 195.87 percent relative improvement in Novel Trajectory Agent IoU (NTA-IoU) — the primary novel-trajectory fidelity metric — over the DriveDreamer4D with PVG baseline. This metric

measures overlap between synthesized and reference agent bounding boxes along novel trajectories; higher scores indicate better fidelity, and the 195.87 percent gain represents a near-tripling of reconstruction accuracy in out-of-distribution regions, establishing a quantitative benchmark for novel-trajectory robustness that subsequent systems can be evaluated against.

DriveDreamer4D [5] complemented this finding by demonstrating that conditioning on high-definition map overlays and three-dimensional agent bounding boxes substantially extends the range of trajectories over which geometric fidelity is maintained. The structured conditioning inputs provide an explicit geometric anchor that prevents scene element drift even when the ego trajectory explores regions of state space not covered by the training data. Together, these two contributions define a testable criterion for world model deployment readiness: a system is suitable for novel-trajectory evaluation only if its NTA-IoU score on a held-out out-of-distribution trajectory set falls within a specified tolerance of its in-distribution performance.

V. Reinforcement Learning Policy Hardening in World-Model Environments

Imitation learning from recorded driver behavior has a well-documented structural limitation for safety-critical policy training: the recorded data is distributed over normal driving situations, and the policy trained on it inherits that distributional bias [17], [19]. A policy that imitates normal driving competently may still lack the specific behavioral patterns required to navigate the tail of safety-critical situations that the recorded drivers rarely encountered.

Reinforcement learning in world-model environments addresses this limitation by replacing the imitation objective with an explicit safety reward signal [15]. The world model serves as the environment: the policy receives observations synthesized by the world model, executes actions, and receives rewards calibrated to safety margins, comfort constraints, and task completion. The training loop runs entirely within the world model, enabling thousands of policy updates per hour without requiring physical vehicle interaction.

The RAD system [23], accepted at NeurIPS 2025, is the first demonstration of end-to-end driving-policy training with reinforcement learning inside a 3D Gaussian-Splatting world model [2]. By building a photorealistic digital twin and designing rewards that penalize collisions, discomfort, and lane-center deviation, RAD's policy achieved a collision rate three times lower than imitation-learning baselines on the same closed-loop benchmark (CARLA Leaderboard 2.0 / Bench2Drive). This result validates the core hypothesis that closed-loop world-model environments can supply the volume and diversity of interactive experience that reinforcement learning needs in safety-critical driving [22]. The three-dimensional Gaussian splatting representation provides the real-time rendering performance required to sustain reinforcement learning training loops at practically useful sample rates, with inference costs substantially lower than diffusion-based rendering [18], enabling the thousands of environment interactions per training episode that reinforcement learning algorithms require.

The reward design challenge is distinct from the rendering challenge: a reward that is too narrowly specified produces a policy that optimizes the specified metric while failing on related safety dimensions; a reward that is too broadly specified produces insufficient training signal for rare-event scenarios. RAD's reward structure — penalizing collision, discomfort, and lane-center deviation jointly — represents a practical resolution of this tension validated against a standardized closed-loop benchmark [24].

VI. Counterfactual Dataset Construction for Safety-Critical Training Data

Near-miss events captured by the fleet represent a uniquely valuable resource for safety-critical training data construction. These events are geometrically grounded in real-world road geometry and agent behavior, they involve the specific physical configurations that are most relevant to collision risk, and they are systematically captured by fleet safety monitoring systems. Their limitation for training data purposes is their rarity and their fixed geometry: each near-miss event occurs once, with a specific set of agent speeds, headings, and gap distances [20], [21].

Counterfactual perturbation within a world model environment converts each near-miss event into a family of calibrated variants. By applying controlled modifications to agent speed profiles, heading

angles, and inter-vehicle gap distances within the world model, a single near-miss event can be expanded into dozens of scenarios spanning a continuous range of collision probabilities [8]. The resulting dataset covers the specific geometric configurations that the near-miss geometry defines, while the perturbation range provides the distributional breadth required for robust policy training.

AD-R1 [11], introduced in late 2025, operationalized this approach through an Impartial World Model framework trained specifically on counterfactual datasets constructed from near-miss recordings perturbed into the collisions and off-road departures they narrowly avoided. Training on these synthesized failure outcomes produces a world model markedly more accurate at predicting failures than one trained on standard fleet logs, which contain almost no actual collisions. When deployed as an internal critic in a closed-loop reinforcement-learning framework, the AD-R1 system produced a substantial reduction in safety violations on challenging scenarios compared with standard world models evaluated on a risk-foreseeing benchmark. Qwen2.5-VL-7B-Instruct was deployed as an automated quality filter [19], scoring each generated counterfactual variant for physical plausibility before admitting it to the training set — an essential step, as perturbations that exceed the physical plausibility envelope produce geometrically implausible scenarios that degrade rather than improve policy performance.

Table 2. Four Applications: Mechanisms and Safety Outcomes

Application	Primary Mechanism	Safety Outcome	System	Reference
City-scale simulation	Joint agent generation; CitySim full-journey protocol	>60 s coherent city-scale simulation; cumulative risk profile across full journey	SceneDiffuser++	[10]
Novel-trajectory assessment	Online restoration; structured conditioning anchor	+195.87% NTA-IoU over DriveDreamer4D/PVG baseline	ReconDreamer	[4]
RL policy hardening	3DGS real-time environment; safety reward design	3× lower collision rate vs. imitation-learning baselines (CARLA Leaderboard 2.0)	RAD	[23]
Counterfactual construction	Near-miss perturbation; Impartial World Model; VLM quality filtering	Substantial drop in safety violations on risk-foreseeing benchmark	AD-R1	[11]

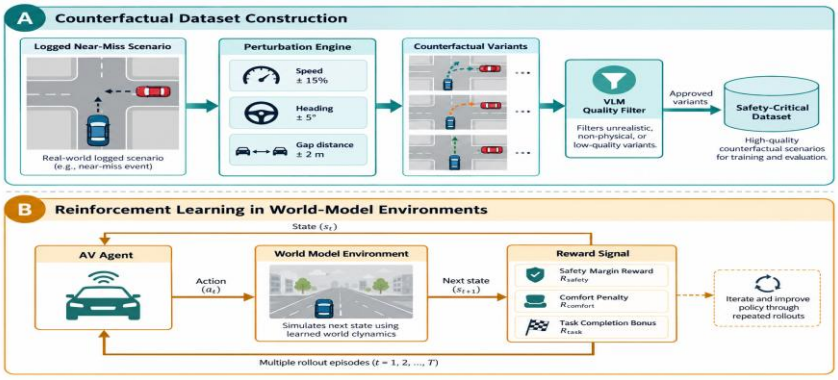


Fig. 2. Integrated workflow for counterfactual dataset construction (top) and reinforcement learning in world-model environments (bottom): near-miss logs are perturbed and quality-filtered to build safety-critical training sets, while the world model serves as the environment for policy optimization through reward-driven rollouts.

VII. Cross-Application Analysis: Three Shared Mechanisms

The four applications documented above share three structural mechanisms that explain their safety impact. Recognizing these mechanisms enables engineering teams to evaluate whether a proposed world model deployment is likely to produce genuine safety improvement or merely improved simulation metrics.

A. Causal Coupling

Each of the four applications depends on a world model that incorporates agent actions into scene state updates at each time step [12], [20]. City-scale simulation requires that background agent trajectories respond coherently to the behavior of the agent under test. Novel-trajectory assessment requires that scene geometry evolve consistently with the ego vehicle's actual commanded trajectory rather than a predicted nominal trajectory. Counterfactual construction requires that perturbed agent behaviors propagate geometrically consistent consequences through the simulated scene. Reinforcement learning requires that the world model environment respond to policy actions with scene transitions that accurately reflect the physical consequences of those actions. A world model that lacks causal coupling in any of these senses will fail to produce genuine safety improvement regardless of its perceptual quality.

B. Distributional Coverage of Rare Events

Each application extends the range of scenarios available for policy evaluation or training beyond what fleet logs can supply [1], [9]. City-scale simulation covers the full joint distribution of agent behavior across complete journeys exceeding sixty seconds. Novel-trajectory assessment systematically probes the boundary of the world model's training distribution, with the 195.87 percent NTA-IoU improvement quantifying the gain in coverage [4]. RAD's reinforcement learning — validated by the threefold collision-rate reduction — enables policy updates in response to synthesized rare-event scenarios that the imitation training distribution does not cover [23]. Counterfactual construction generates calibrated variants of rare near-miss geometries, building a dataset that covers collision-probability ranges not present in the fleet log [11]. The common thread is deliberate scenario generation targeting the tail of the safety-relevant distribution, rather than sampling from the empirical fleet log distribution.

C. Reward-Aligned Policy Optimization

The reinforcement learning application makes explicit what the other three applications imply: the training objective must be aligned with safety margins rather than with perceptual realism or imitation fidelity [15], [17]. City-scale simulation provides the scenario scope for evaluating policies against cumulative safety metrics [10]. Novel-trajectory assessment provides the diagnostic signal for identifying policy failure modes at distribution boundaries [4]. Counterfactual construction provides the training data distribution required to harden the policy against the specific geometric configurations most associated with collision risk [11]. All three mechanisms are prerequisites for reward-aligned policy optimization; none is sufficient on its own.

Table 3. Three Cross-Cutting Mechanisms: Application Dependency Matrix

Application	Causal Coupling	Distributional Coverage	Reward Alignment	Deployment Readiness
City-scale simulation	Required	Primary	Supporting	High
Novel-trajectory assessment	Required	Diagnostic	Supporting	Moderate-High
RL policy hardening	Required	Required	Primary	Moderate
Counterfactual construction	Required	Primary	Required	Moderate

VIII. Deployment Timing Framework

The four applications do not have uniform deployment prerequisites. City-scale simulation requires a world model with joint agent generation capability and sufficient temporal scope to sustain full-journey evaluation beyond sixty seconds [10]; these are demanding architectural requirements that only a subset of currently available systems satisfy. Novel-trajectory assessment can be layered onto any world model deployment as a diagnostic tool once a baseline closed-loop evaluation protocol is in place [4], [5]. Counterfactual dataset construction requires a world model with structured scene conditioning capability sufficient to propagate perturbations geometrically consistently, plus a calibrated quality filter [11]. Reinforcement learning in world-model environments additionally requires real-time or near-real-time inference performance, which restricts the compatible architecture to occupancy-centric and Gaussian splatting hybrid systems [2], [3], [23].

The deployment sequence that minimizes risk while maximizing early safety impact is as follows. Novel-trajectory assessment is the appropriate first application because it provides immediate diagnostic value for any existing world model deployment without requiring additional architectural investment [4]. City-scale simulation with full journey evaluation is the appropriate second application for teams whose world model infrastructure already supports joint agent generation and sixty-second coherent rollouts [10]. Counterfactual dataset construction is the appropriate third application, requiring structured conditioning capability and a deployed quality filter [11]. Reinforcement learning policy hardening is the appropriate fourth application, requiring all of the above plus real-time inference performance [23], [24].

Organizations that attempt to deploy reinforcement learning policy hardening without first establishing closed-loop evaluation protocols and counterfactual dataset construction capability will find that the reward signal is insufficiently grounded in real-world safety-critical geometry to produce robust policy improvement [15], [22]. The deployment timing framework is therefore not merely a practical recommendation but a prerequisite dependency structure: each application provides the safety grounding that makes the next application's training signal meaningful.

IX. Discussion

The transition from open-loop to closed-loop world model simulation represents a structural shift in how autonomous vehicle safety validation is conducted, not merely an incremental improvement in simulation fidelity [13], [14]. Open-loop simulation answers the question of whether a world model produces plausible-looking outputs. Closed-loop simulation with causal coupling answers the question of whether a policy trained on those outputs behaves safely when the consequences of its actions propagate through the simulated environment. The distinction is not subtle, and the engineering investment required to close the gap between these two questions is substantial.

The three mechanisms identified in Section VII — causal coupling, distributional coverage, and reward alignment — provide a diagnostic framework for evaluating any proposed world model safety application. An application that cannot demonstrate all three mechanisms is unlikely to produce genuine safety improvement regardless of its simulation quality metrics. This framework is intentionally technology-agnostic: it applies equally to diffusion-based [6], [18], occupancy-centric [3], and Gaussian splatting architectures [2], [23], and it will remain applicable as these architectures evolve.

The quantitative results documented here — sixty-second coherent city-scale rollouts from SceneDiffuser++ [10], a 195.87 percent NTA-IoU improvement from ReconDreamer [4], a threefold collision-rate reduction from RAD [23], and a substantial safety-violation reduction from AD-R1 [11] — collectively establish that closed-loop generative simulation is no longer a research curiosity but a production-relevant engineering capability. The primary unresolved barriers are computational cost for diffusion-based systems, reward specification depth for reinforcement learning applications, and automated quality assurance at the scale of continuous fleet log ingestion. Each of these barriers is the subject of active research, and the trajectory of progress over the past two years suggests that production-grade deployment is a near-term rather than speculative prospect for well-resourced development programs [7], [8].

X. Conclusion

Closed-loop generative world model simulation has moved from a research curiosity to a production-relevant engineering capability within the span of two years. The four applications documented in this article — city-scale traffic simulation sustaining coherent rollouts beyond sixty seconds [10], novel-trajectory assessment with a documented 195.87 percent improvement in NTA-IoU fidelity [4], reinforcement learning policy hardening achieving a threefold collision-rate reduction [23], and counterfactual dataset construction producing a substantial drop in safety violations [11] — each address a distinct structural limitation of existing autonomous vehicle safety validation methodologies, and each has produced documented safety outcomes in peer-reviewed evaluations.

The three cross-cutting mechanisms identified through cross-application analysis — causal coupling, distributional coverage of rare events, and reward-aligned policy optimization — provide a principled basis for evaluating whether a proposed world model safety application is likely to produce genuine fleet safety improvement. Safety engineering teams can apply these criteria to evaluate the readiness of their world model infrastructure for each application tier, and the deployment timing framework developed in Section VIII maps infrastructure readiness to application priority.

Future work should prioritize three areas: standardized closed-loop benchmark protocols that enable cross-system comparison on safety-relevant scenario sets [1], [9], automated reward specification methods that reduce the expert engineering time required to deploy reinforcement learning in world-model environments [15], [21], and efficient world model architectures that eliminate the computational cost differential between diffusion-quality synthesis and real-time simulation performance [7], [8], [18].

References

[1] H. Caesar et al., "nuScenes: A multimodal dataset for autonomous driving," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2020, pp. 11621–11631. [Online]. Available: https://arxiv.org/abs/1903.11027

[2] S. Zhou et al., "GaussianWorld: Gaussian world model for streaming 3D occupancy prediction," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2025. [Online]. Available: https://arxiv.org/abs/2412.10373

[3] Bohan Li et al., "UniScene: Unified occupancy-centric driving scene generation," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2025. [Online]. Available: https://arxiv.org/abs/2412.05435

[4] Ch. Ni et al., "ReconDreamer: Crafting world models for driving scene reconstruction via online restoration," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2025. [Online]. Available: https://cvpr.thecvf.com/virtual/2025/poster/32804

[5] Guosheng Zhao et al., "DriveDreamer4D: World models are effective data machines for 4D driving scene representation," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2025. [Online]. Available: https://arxiv.org/abs/2410.13571

[6] Guosheng Zhao et al., "DriveDreamer-2: LLM-enhanced world models for diverse driving video generation," arXiv preprint arXiv:2403.06845, 2024. [Online]. Available: https://arxiv.org/abs/2403.06845

[7] Niket Agarwal et al., "Cosmos World Foundation Model Platform for Physical AI," 2025. [Online]. Available: https://research.nvidia.com/labs/toronto-ai/cosmos-av/

[8] Santos, C. P., Martino, L., Gonzalez, R. S., & Moran, J. J. D. (2026). Technological Gaps and Security: Challenges for Law Enforcement and Public Safety. In S. Cea & T. Aznar (Eds.), AI Influence on Governance and Law in the Digital Age (pp. 169-192). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-4480-5.ch007

[9] X. Ren et al., "Cosmos-Drive-Dreams: Scalable synthetic driving data generation with world foundation models," arXiv preprint arXiv:2506.09042, 2025. [Online]. Available: https://arxiv.org/abs/2501.03575

[10] P. Sun et al., "Scalability in perception for autonomous driving: Waymo open dataset," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2020, pp. 2446–2454. [Online]. Available: https://arxiv.org/abs/1912.04838

[11] S. Tan et al., "SceneDiffuser++: City-Scale Traffic Simulation via a Generative World

Model," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2025, pp. 1570–1580. [Online]. Available: https://arxiv.org/abs/2506.21976

[12] Tianyi Yan et al., "AD-R1: Closed-Loop Reinforcement Learning for End-to-End Autonomous Driving with Impartial World Models," arXiv preprint arXiv:2511.20325, 2025. [Online]. Available: https://arxiv.org/abs/2511.20325

[13] Ze Yang et al., "UniSim: A neural closed-loop sensor simulator," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2023, pp. 1389–1399.

[14] Xiaofeng Wang et al., "DriveDreamer: Towards real-world-driven world models for autonomous driving," in Proc. Eur. Conf. Comput. Vis. (ECCV), 2024. [Online]. Available: https://arxiv.org/abs/2309.09777

[15] A. Hu et al., "GAIA-1: A generative world model for autonomous driving," arXiv preprint arXiv:2309.17080, 2023. [Online]. Available: https://arxiv.org/abs/2309.17080

[16] D. Hafner et al., "Dreamer: Learning latent dynamics for planning from pixels," in Proc. Int. Conf. Learn. Represent. (ICLR), 2020. [Online]. Available: https://arxiv.org/abs/1912.01603

[17] S. Suo et al., "TrafficSim: Learning to simulate realistic multi-agent behaviors," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2021, pp. 10973–10982. [Online]. Available: https://openaccess.thecvf.com/content/CVPR2021/papers/Suo_TrafficSim_Learning_To_Simulate_Realistic_Multi-Agent_Behaviors_CVPR_2021_paper.pdf

[18] Y. Hu et al., "Planning-oriented autonomous driving," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2023, pp. 17853–17862. [Online]. Available: https://arxiv.org/abs/2212.10156

[19] J. Ho et al., "Denoising diffusion probabilistic models," in Proc. Advances in Neural Inf. Process. Syst. (NeurIPS), 2020, pp. 6840–6851. [Online]. Available: https://arxiv.org/abs/2006.11239

[20] J. Mao et al., "GPT-Driver: Learning to drive with GPT," arXiv preprint arXiv:2310.01415, 2023. [Online]. Available: https://arxiv.org/abs/2310.01415

[21] M. Yang et al., "CausalVAE: Disentangled representation learning via neural structural causal models," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2021, pp. 9593–9602.

[22] Z. Li et al., "BEVFormer: Learning bird's-eye-view representation from multi-camera images via spatiotemporal transformers," in Proc. Eur. Conf. Comput. Vis. (ECCV), 2022. [Online]. Available: https://arxiv.org/abs/2203.17270

[23] Shenyuan Gao et al., "Vista: A generalizable driving world model with high fidelity and versatile controllability," in Proc. Advances in Neural Inf. Process. Syst. (NeurIPS), 2024.

[24] Hao Gao et al., "RAD: Training an end-to-end driving policy via large-scale 3DGS-based reinforcement learning," in Proc. Advances in Neural Inf. Process. Syst. (NeurIPS), 2025. [Online]. Available: https://arxiv.org/abs/2502.13144

[25] Zhenjie Yang et al., "Raw2Drive: Reinforcement learning with aligned world models for end-to-end autonomous driving," arXiv preprint arXiv:2505.16394, 2025. [Online]. Available: https://arxiv.org/abs/2505.16394

Declarations

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest: The author declares no conflict of interest.

Data Availability: This article does not report new empirical data. All referenced datasets are publicly available as cited.

Ethics Approval: Not applicable.

Author Contributions: Tamoghna Das conceived the study, conducted the literature review, developed the cross-application analysis and deployment framework, and wrote the manuscript in its entirety.

AI Tool Disclosure: The author used AI-assisted writing tools for language refinement and editorial polishing of this manuscript. All scholarly content, research analysis, literature synthesis, and intellectual contributions are the author's own. The author accepts full responsibility for the accuracy and integrity of all content.