# Grounded Anomalies: Towards Causally Grounded Kinematic Anomaly Generation

Joon-Seok Kim
joonseok.kim@emory.edu
Emory University
Atlanta, GA, USA

Andreas Züfle
azufle@emory.edu
Emory University
Atlanta, GA, USA

## Abstract

Anomaly detection plays a critical role in mobility systems by identifying unexpected behaviors that deviate from normative patterns. It supports essential applications such as safety assurance, security monitoring, and post-incident analysis. Despite its importance, acquiring high-quality anomaly data remains a significant challenge due to the rarity of anomalous events and the difficulty of accurate annotation thereof. Due to this lack of ground truth data annotations, prior research predominantly focuses on creating synthetic anomalies by augmenting trajectories with anomalous behavior such as excessive speed or sharp turns. But we argue that detecting anomalies in a vacuum is not a useful task. What we really are interested in is the cause that leads to anomalous behavior: The cause for which the anomalous behavior is only the symptom. For example, a traffic accident caused by an adverse health event may be preceded by more subtle deviations in motions patterns such as unusual speed, acceleration, or turning behavior. These subtle deviations in motion patterns, often resulting from cognitive impairments, environmental stressors, or degraded motor control, can precede more severe events and offer valuable insights into human behavior. But generating and finding causally grounded kinematic anomalies has received limited attention.

In this paper, we address this gap by introducing a generative modeling approach for simulating kinematic anomalies in ground-based mobility contexts, with a primary focus on driving and extensibility to walking and bicycling. We analyze key factors and configurations that influence the manifestation of anomalous behaviors and propose an algebraic framework to modularly generate such behaviors. This approach facilitates synthetic data generation for anomaly detection models and supports scenario design for behavior analysis and safety evaluation.

## CCS Concepts

• **Computing methodologies → Modeling and simulation**; • **Information systems → Spatial-temporal systems**; • **Applied computing → Transportation**.

## Keywords

kinematic anomaly, anomaly detection, mobility, simulation, data generation, causality


*GeoAnomalies '25, Minneapolis, MN, USA*

ACM ISBN 979-8-4007-2260-8/2025/11
https://doi.org/10.1145/3764914.3770591

**ACM Reference Format:**
Joon-Seok Kim and Andreas Züfle. 2025. Grounded Anomalies: Towards Causally Grounded Kinematic Anomaly Generation. In *The 2nd ACM SIGSPATIAL International Workshop on Geospatial Anomaly Detection (GeoAnomalies '25), November 3–6, 2025, Minneapolis, MN, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3764914.3770591

## 1 Motivation: Towards Explaining, Predicting, and Prescribing Anomalies

Detecting anomalies, as an isolated task by itself, might not be very useful. However, it can become highly valuable if the detected anomalies can be used by additional inference tasks, such as: (1) understanding what causes the anomalies and why they occur, (2) predicting and preparing for the future occurrences of anomalies, and (3) addressing the root causes of the anomalies to prescribe changes in the real world to prevent future anomalies from occurring. As an example, imagine traveling back in time year 1854, to the Soho district (now part of the City of Westminster) in London, UK. At that time, Soho was plagued by a severe anomaly: a cholera outbreak. The outbreak was short-lived but very intense: over 500 people died in just about 10 days. Figure 1 shows the famous map of John Snow [44] published in 1854, which pioneered the field of spatial epidemiology. The map illustrates the road network of Soho annotated with observed cholera cases. Now, imagine a similar outbreak occurring today, with modern sensing devices capturing the location of movement of individuals. How could trajectory anomaly detection be used to respond to a cholera outbreak in 2025?

*Task 0: Anomaly Detection.* By analyzing and mining human trajectories, we should be able to detect individual trajectories that exhibit changes in mobility behavior or stop moving entirely. However, this knowledge, by itself in a vacuum, is not actionable. Without knowing what the cause of the behavior change is, it could simply be a stormy day causing people to stay indoors.

*Task 1: Anomaly Explanation.* Given the results from Task 0, we can start investigating. Public health officials may be able to explain that the change in mobility is due to cholera infection, and that the complete cessation of movement in some trajectories is likely due to death. With this explanation, the detected anomalies can be used to inform first responders, who may then visit these locations to check on potentially deceased individuals. While this result has some utility (e.g., removing the bodies of individuals who lived alone), this would not likely save any lives, since cholera is not transmitted between people or from dead bodies.

*Task 2: Anomaly Prediction.* Given the results from Task 0 and the explanation from Task 1, we can then use predictive modeling to predict the number and locations of future individuals becoming

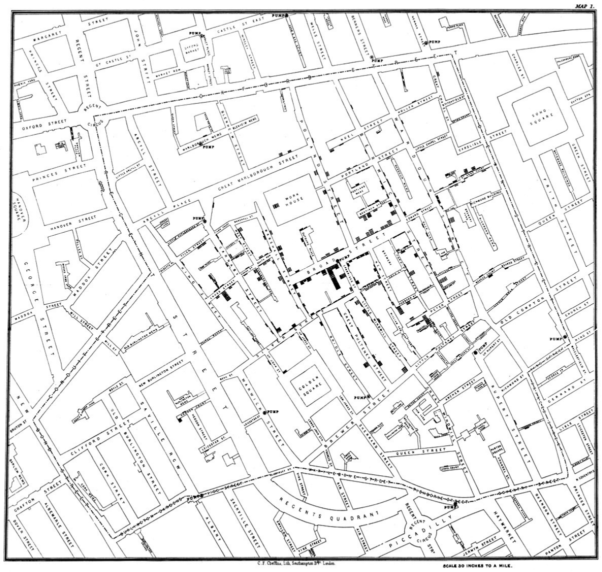

**Figure 1: John Snow's 1854 map of the Soho Cholera Outbreak [44] (image public domain)**

infected by and succumbing to cholera. When used for predictive analytics, anomaly detection begins to have life-saving potential: we may be able to evacuate areas with high predicted risk or protect vulnerable populations.

*Task 3: Anomaly Prescription.* If we can detect anomalies (Task 0), explain their causes (Task 1), leverage this understanding for predicting future anomalies (Task 2), the final and most impactful task is to take actions to improve future outcomes: to prescribe changes to the real world that prevent future anomalies. In the case of John Snow, he correctly identified the Broad Street water pump as the source of cholera infections, shut it down, and thereby stopped the outbreak, saving lives and preventing further infections.

Following this example of the 1854 Cholera Outbreak, we argue that anomaly detection is merely the task of identifying symptoms. The real impact stems from understanding these symptoms and prescribing interventions to address their underlying causes. However, many existing works stop at Task 0, and structure their research in a way that cannot progress beyond it: often by using data in which anomalies are inserted arbitrarily, without any causal process.

For example, recent work in the machine learning community on kinematic trajectory anomaly detection generates anomalies into datasets arbitrarily. Many recent studies [15, 18, 28, 32, 50, 54, 56] augment data by injecting so-called *Detour Anomalies* (where a trajectory is altered to take a longer route than observed) and *Switch Anomalies* (where segments of a trajectory are swapped with segments from another trajectory) at arbitrary times and locations. In addition, the recent work Shao et al. [42] introduced so-called *Time Anomalies* (where the speed of a trajectory segment is increased or decreased) and *Loop Anomalies* (where reflexive detours from one location back to itself are added). But all of these anomalies have in common that they have no causal grounding: they are selected arbitrarily. In the context of the Soho Cholera Outbreak, this is like creating an anomalous trajectory dataset by taking an existing (normal) trajectory dataset and selecting, arbitrarily at uniform random, trajectories to become dead (i.e., no longer moving). While such a dataset still allows to perform Task 0 (detecting the affected trajectories), it is not possible to perform Task 1. That is because the affected agents lack any causal grounding; their status is determined by entropy, not by a meaningful cause. Attempting to explain the cause of detected anomalies in such datasets is futile, as no cause exists. Consequently, Task 2 (prediction) is also futile, since the anomalies are selected uniformly at random, resulting in unexplainable variance. By definition, such variance cannot be explained or predicted. Without any causal basis for these inserted anomalies, Task 3 of prescribing actions to prevent future anomalies, is ill-defined.

## 2 Introduction: Towards Causally Grounded Kinematic Anomalies

Understanding and modeling anomalous behaviors mobility systems is essential for advancing safety and robustness in autonomous transportation systems [5] and has applications in monitoring elders, and tracking the spread of infectious diseases [57]. Anomalies, which are behaviors that deviate from typical patterns, can signal potential safety hazards, cognitive impairments, or environmental stressors. Kinematic anomalies manifest in the change of the mechanical motion of an object, such as anomalous (high/low) speed, anomalous acceleration (deceleration), or anomalous (fast or slow in a curve) changes of direction. Such kinematic anomalies may have a multitude of causal reasons, such as a change of the environment (such as fog leading to many vehicles driving slower than normal), a medical hazard (such as a cramp or a stroke of a driver leading to loss of control), or emotional distress (such as a verbal argument with a passenger leading to rapid acceleration). The state-of-the-art detects kinematic anomalies in hindsight: To determine the cause of a traffic accident.

Our vision is to advance kinematic anomaly detection to the next level and to monitor live traffic. Detected anomalies can then trigger real-time interventions, such as alerting the anomalous driver, warning nearby other drivers, or calling an ambulance even before an accident happens. This may give first-responders a potentially lifesaving head start. For example, an anomalous lack of directional change while approaching a curve could prompt a driver's assistant system to ask the driver, "Are you OK?". The driver might respond, "It's suddenly so foggy!" or "H-h-he-he-help" with the latter response automatically triggering a call for emergency services.

Towards this vision, however, one of the major challenges is the scarcity of high-quality, well-annotated anomaly data, particularly for fine-grained kinematic deviations that often precede overt failures. The goal of this vision paper is to chart a path forward towards simulating kinematic anomalies and their causal reasons to create large datasets of mobility trajectories with anomaly labels [60]. Such a dataset would allow us to investigate what types of anomalies can be detected rapidly enough to allow interventions and whether such a system detection system may create an overwhelming number of false positives.

To simulate kinematic anomalies, we acknowledge that driving behavior is influenced by a complex interplay of cognitive, environmental, demographic, and contextual factors. Human decision-making in dynamic road environments is inherently variable and susceptible to both systematic biases (such as age-related perceptual decline) and spontaneous disruptions like emotional stress or distraction. Despite growing research in driver behavior modeling, current systems often treat anomalous behavior as an outlier, rather than as a structured deviation that can be systematically analyzed or generated. Recent studies have advanced our understanding of specific driving impairments and the diversity of decision-making strategies. For instance, age has been linked to both increased self-regulation and variability in cognitive function [4, 21, 33]. Distraction and cognitive workload have been shown to significantly degrade driving performance [16, 27, 36], while emotional states such as anger or stress are correlated with risky maneuvers and reduced situational awareness [17, 19]. Environmental factors such as weather conditions [11, 38], visibility [53], road signage [13], and cultural differences [45] further complicate behavioral modeling.

In parallel, the emergence of intelligent transportation systems has driven interest in data-driven approaches for anomaly detection [12, 39] and driver profiling [30, 46]. However, most existing methods focus on classification or detection, limiting their utility in proactive simulation, scenario generation, and model robustness testing. More notably, the majority of research targets high-level anomalous events (e.g., accidents or lane violations) while subtler kinematic anomalies, which may serve as early warning signals, remain underexplored.

In this paper, we address this gap by focusing on the simulation and generation of fine-grained kinematic anomalies in ground mobility contexts. We propose a generative modeling framework that incorporates behavioral, demographic, and environmental decision-making factors. Unlike prior work, our approach formalizes the process of generating structured deviations from normative mobility behavior, allowing for the synthesis of realistic anomalous scenarios.

While this work emphasizes driving due to its complexity and rich behavioral variability, the proposed framework is designed to be modality-agnostic. Other ground-based mobility forms such as bicycling and walking also exhibit kinematic anomalies, particularly under impaired or distracted conditions. By extending our model to these modalities, we provide a broader foundation for studying human mobility behavior in heterogeneous traffic environments. Through this lens, we aim to contribute to the development of safer, more adaptive mobility systems by providing a systematic method for generating and analyzing kinematic anomalies. Our work facilitates not only improved anomaly detection and classification but also supports the design of behaviorally diverse simulation environments for training and evaluation.

## 3 Related Work

We survey the literature on kinematic anomalies and the various factors that influence data-driven modeling of anomalous behaviors in mobility, with particular attention to driving contexts. Our review spans work on kinematic definitions, decision-making factors, and recent advances in generative modeling and intelligent systems.

### 3.1 Kinematic Anomalies

Kinematic anomalies refer to deviations in motion patterns such as speed, acceleration, and trajectory curvature that diverge from expected behavior. Kennedy and Züfle [22] formalize this by introducing the concept of a "kinematic profile," a personalized statistical signature of motion. Deviations from this profile, which is measured through features like velocity distributions and acceleration variance, are treated as indicators of anomalous behavior. Similarly, Xiao et al. [55] analyze human movement in video sequences by extracting explicit kinematic features such as stride length and limb displacement. Their framework detects subtle motion anomalies in real time, illustrating the utility of fine-grained kinematic monitoring in broader behavior analysis. Following these initial studies, the machine learning community has recently started to tackle the challenge of detecting kinematic anomalies: A very recently published ACM KDD Paper [42] uses a Large Language Model-based approach to detect kinematic anomalies such as fast turns or slow driving.

It may be argued that a short-coming of these existing works is that they specializes on very specific types of known and predefined anomalies. Yet, the challenge of (unsupervised) outlier detection in contrast to (supervised) classification is that there is no training data and there are no examples of what an anomaly looks like. Thus, it remains unclear how these existing works may generalize to other anomalies than the ones they define. Our work aims at filling this gap, by providing a framework that can generate broad variety of kinematic anomalies would allow the anomaly detection a more comprehensive evaluation of what they can detect.

### 3.2 Decision-Making Factors in Driving

Modeling anomalous driving behavior requires accounting for the multifaceted nature of human decision-making, influenced by cognitive, demographic, emotional, environmental, and contextual elements. We categorize the literature into four key themes.

*Cognitive and Demographic Factors.* Age, perception, and cognitive capacity are foundational variables in driver behavior modeling. Bernstein et al. [4] and Menze et al. [33] report that older drivers often engage in more self-regulation, though executive function is a stronger predictor of risky decisions. Joshi et al. [21] corroborate these findings using real-time vehicle data, showing divergent control patterns across age groups. Zhang and Liu [58] highlight the heterogeneity of crash risk among older drivers. Meanwhile, Molnar et al. [34] and Watson-Brown et al. [52] identify inconsistencies between perceived and actual driving safety, complicating the modeling of subjective risk.

*Distraction, Stress, and Emotional States.* Distraction is a critical contributor to unsafe behavior. Garcia-Constantino et al. [16] employ time-series analysis to quantify how cognitive distraction affects vehicle control. EEG-based studies like Li et al. [27] offer physiological validation for distraction detection. Nakano and Chakraborty [7, 35, 36] demonstrate the potential of deep learning for real-time detection of driver awareness based on telemetry data.

Emotional influences, particularly anger and stress, also impair driving behavior. Brewer [6] and González-Iglesias et al. [17] associate emotional dysregulation with higher traffic violations, with gender-specific nuances. Hill and Boyle [19] link driver stress to both maneuver complexity and road conditions. Singh and Kathuria [43] provide a comprehensive review of these psychological impacts in naturalistic driving studies.

*Behavior Modeling and Driver Profiling.* Capturing personalized driving styles is vital for both anomaly detection and simulation. Tselentis and Papadimitriou [46] identify challenges in driver profiling, emphasizing the need for context-aware modeling frameworks. Liao et al. [30] provide an overview of datasets and methodologies for personalization in driving behavior modeling, while their generative work on human mobility patterns [29] aligns with the goals of anomaly synthesis in our study.

Deep learning has shown promise in this domain. Praharsha and Poulose [39] achieve high accuracy in distraction classification using attention-based architectures (CBAM-VGG16). Fan et al. [12] use hybrid models to detect anomalous lane changes, highlighting the capacity of neural models to capture subtle behavioral deviations.

*Multimodal and Intelligent Systems.* Recent advances integrate behavioral analysis with multimodal sensing and contextual awareness. Wang et al. [51] propose AccidentGPT, a large multimodal model leveraging vehicle-to-everything (V2X) data for accident prediction. Varnosfaderani et al. [47] review unobtrusive biosensing systems for real-time cognitive and emotional monitoring. Amiri et al. [3] emphasize autonomous learning strategies for modeling complex behavioral interactions in interconnected environments.

## 3.3 Environmental and Contextual Factors

Environmental and infrastructure variables significantly shape driving behavior and may induce or exacerbate anomalies. Numerous studies have demonstrated that adverse weather, such as fog or ice, leads to impaired maneuvering and elevated crash risk [8, 11, 37, 38, 41]. Road infrastructure quality, including signage, markings, and pavement, directly influences cognitive load and driver gaze behavior [13, 26]. Visibility, gender, and cultural norms further modulate the impact of environmental factors, as shown by Wei et al. [53] and Taourarti et al. [45]. To summarize, while current research has made strides in detection and profiling, limited attention has been paid to the structured generation of fine-grained kinematic anomalies. This motivates the present study, which seeks to formalize and simulate such behaviors, offering a foundation for more realistic testing, analysis, and training of intelligent mobility systems.

# 4 Kinematic Anomalies

An anomaly is defined as a deviation from expected or normal behavior [14]. In the context of mobility, the definition of "normal" is inherently subjective and depends on modeling assumptions and context. This work focuses specifically on characterizing such behavior through a *kinematic* lens, centering on measurable motion variables.

Our core research question is: *How can we systematically generate kinematic anomalies within a simulation environment using observable features, while grounding them in plausible, semantically meaningful deviations?*

In this paper, we restrict our focus to ground-based mobile entities, such as pedestrians and vehicles, and exclude other transportation modalities like aerial or aquatic systems for simplicity. Table 1 outlines the scope and boundaries of this study. Human mobility anomalies in patterns of life [2] or gait [55] models are not considered in this study. Within simulations, sensor errors (e.g., GPS inaccuracies) are considered plausible and relevant for autonomous vehicle decision-making. However, any assumptions regarding observation errors in output data generated externally are beyond the scope of this study. This ensures that users can access ground truth data, including accurate kinematic variables.

**Table 1: Scope of this study**

| Type | Scope | Out of Scope |
|---|---|---|
| Mode | Pedestrian, Ground vehicle | Watercraft, Aerial vehicle |
| Model | Kinematic behavior | Gait, Patterns of life |
| Error | Sensor errors within simulation | Observation noise in recorded outputs |

## 4.1 Control-Theoretic View of Mobility

We adopt a control-theoretic perspective to structure the components involved in mobility. In classical control systems, behavior is governed by three core elements: sensors, controllers, and actuators. This abstraction is foundational in modern vehicle dynamics and autonomous mobility design [31, 40]. Table 2 illustrates how these components map to the mobility domain.

A sensor may refer to human vision or machine perception systems such as cameras, LiDAR, radar, and GPS, which observe the surrounding environment. The controller represents the decision-making unit and this could be a human driver, an autonomous agent, or an AI assistant responsible for interpreting sensor data and generating appropriate actions. The actuator is the physical mechanism that executes these actions, such as throttle, brake, or steering systems–or, in the case of walking, human limbs.

Mobility behavior can be viewed as the process of minimizing deviation from a desired reference state, such as target speed, heading, or lane position, while responding to external disturbances like adverse weather, road conditions, or dynamic traffic contexts.

**Table 2: Mapping between control theory and mobility**

| Control Theory | Mobility Analogue |
|---|---|
| Sensor | Human vision, camera, LiDAR, GPS, radar |
| Controller | Human driver, autonomous agent, AI assistant |
| Reference State | Position, velocity, heading, relative distance |
| Actuator | Throttle, brake, steering, limbs |
| Disturbance | Weather, road condition, traffic context |

## 4.2 Modeling Kinematic Anomalies

We model the behavior of a mobile agent using a sequence of observable state variables. These variables are categorized into two primary types:

- **Core kinematic variables:** Directly measurable motion quantities such as position ($p$), velocity ($v$), acceleration ($a$), and jerk ($j$). These are typically represented as vectors in 2D or 3D space, capturing both magnitude and direction of movement.
- **Contextual variables:** Additional descriptors that enrich behavioral interpretation, including heading ($\theta$), angular velocity ($\omega$), road geometry, proximity to other agents, and interactions.

We denote the complete observable state at each timestep as $X_t$, where $X = \{X_1, X_2, \ldots, X_T\}$ is the sequence of observations over

**Table 3: Summary of Kinematic Anomaly Categories and Modal Examples**

| Category | Modality | Example Anomalies |
|---|---|---|
| **Core Kinematic** | Vehicle | Off-lane driving, sudden stops, harsh braking, throttle-brake oscillation |
| | Pedestrian | Unexpected stops in crosswalks, sprinting in crowded zones |
| | Cyclist | High-speed riding in pedestrian zones, sudden deceleration |
| | Robot | Jerky movement in corridors, delayed acceleration or stopping |
| **Contextual** | Vehicle | Illegal U-turns, failure to yield, wrong-way driving |
| | Pedestrian | Ignoring traffic lights, entering vehicle lanes |
| | Cyclist | Riding against traffic on one-way roads, failure to stop at signs |
| | Robot | Entering human-only areas, navigating against flow in shared spaces |
| **Compound (Multi-factor)** | Vehicle | Driving in reverse on active road (heading + road semantics) |
| | Pedestrian | Diagonal sprint across intersection on red (velocity + violation) |
| | Cyclist | Speeding through crowded intersection against signal |
| | Robot | Backward entry into a congested elevator (heading + proximity + context) |

time. These observations are assumed to be available from sensors or simulation outputs.

To support both generative modeling and detection, we conceptually distinguish between:

- **Internal (latent) state** $Z$: Includes unobservable or inferred factors such as intentions, attention, control policies, and hidden environmental parameters.
- **Observable state** $X$: Includes measurable motion and context variables derived from sensor data or simulation.

We note that anomaly detection systems operate over $X$, while anomaly generation in simulation may additionally utilize $Z$ to simulate plausible but unexpected behaviors.

We define a kinematic anomaly probabilistically as follows:

**Definition 1** (Kinematic Anomaly)**.** A kinematic anomaly is a statistically significant deviation from expected patterns over observable state variables:

$$P(Y = 1 \mid X_{1:T}) \gg \mathbb{E}_{\text{normal}}[P(Y = 1 \mid X_{1:T})],$$

where $Y = 1$ indicates an anomalous sequence, $X_{1:T}$ is the sequence of observed states over time, and $\mathbb{E}_{\text{normal}}[\cdot]$ denotes the expected anomaly probability under a distribution of normal behavior.

This definition enables a flexible, data-driven approach to anomaly detection. Instead of relying on static thresholds (e.g., "velocity > 80 mph"), it allows probabilistic models to identify rare or unexpected patterns within high-dimensional behavior distributions.

In practice, a dynamic detection criterion may be used:

$$P(Y = 1 \mid X_{1:T}) > \mathbb{E}_{\text{normal}}[P(Y = 1 \mid X_{1:T})] + k\sigma,$$

where $k$ is a tunable sensitivity parameter, and $\sigma$ is the standard deviation of the anomaly score under the normal distribution.

While derived features such as time-to-collision or relative speed can enhance detection, they are not strictly required by this formalism, which focuses on deviations from learned distributions in observable states.

## 4.3 Kinematic Anomaly Categories and Examples

We categorize kinematic anomalies into three main types: **core kinematic anomalies**, which involve deviations in directly measurable motion variables (e.g., position, velocity, acceleration, jerk); **contextual anomalies**, which arise from violations of semantic rules, traffic norms, or environmental interactions; and **compound anomalies**, which result from joint deviations across multiple variables or factors. These categories, outlined in Table 3, apply across various modalities including vehicles, pedestrians, cyclists, and robots.

*Core Kinematic Anomalies.* These anomalies arise from deviations in low-level motion dynamics. Since they involve measurable signals like position, speed, and acceleration, they are well-suited for detection using sensor-based systems and signal processing techniques. Temporal consistency is important: a sudden velocity spike might be benign unless it persists or co-occurs with other anomalies over time.

*Contextual Anomalies.* These are defined relative to scene semantics and interaction norms. Their detection requires reasoning over spatial, temporal, and social context (e.g., map data, nearby agent behavior, or traffic rules). For example, a pedestrian jaywalking during a red light may only be anomalous when contextualized within the current phase of the signal and presence of traffic.

*Compound Anomalies.* These combine deviations across both core and contextual dimensions, typically manifesting over extended time windows. Their detection benefits from sequence models (e.g., RNNs, Transformers) that can capture long-term dependencies. Compound anomalies often indicate higher risk levels, as they reflect deeper failures in decision-making, perception, or intent alignment.

*Temporal Implications.* Temporal modeling plays a key role in anomaly detection. Anomalies may:

- **Emerge gradually:** e.g., drift off the lane over time.
- **Occur abruptly:** e.g., sudden stop in a fast-moving lane.
- **Be transient or persistent:** transient anomalies may self-correct; persistent ones indicate systematic failure.

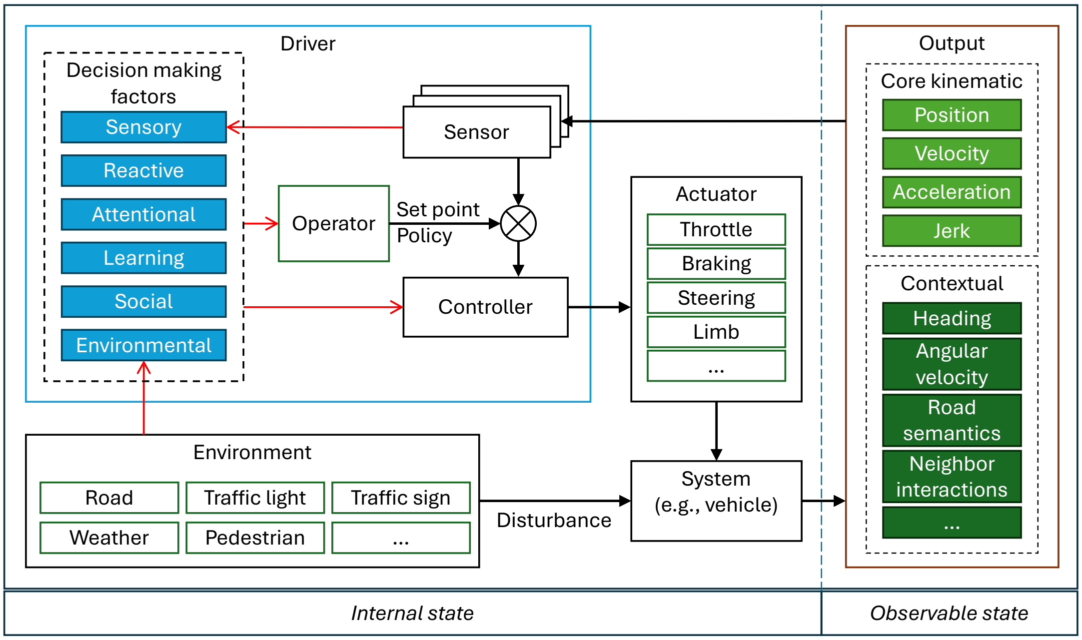


**Figure 2: A schematic of the generative framework for kinematic anomalies.**

For effective detection, systems must not only assess the instantaneous state but also reason over temporal dynamics, historical trends, and predictive uncertainty.

*Detection Implications.* Different anomaly categories require different detection approaches:

- **Core anomalies** can be flagged using statistical modeling, thresholding, or trajectory forecasting [22].
- **Contextual anomalies** require fusion of map knowledge, intent inference, and interaction modeling [20].
- **Compound anomalies** benefit from hybrid methods that combine rule-based priors, simulation-based expectation, and deep sequence modeling [9, 49].

Ultimately, the design of detection models should align with the anomaly types and the observability of variables involved.

## 5 Generative Framework for Kinematic Anomalies

We design a generative model to simulate anomalies through systematic manipulation of internal states and external conditions. Figure 2 shows a schematic of the generative framework for kinematic anomalies. At the core of the model is a **driver module**, which can represent a human, an autonomous agent, or a hybrid AI system. The driver comprises sensors, operators, and a controller, and interacts with a set of decision-making factors, grouped as follows:

- **Sensory:** Inputs that form the driver's perception of the environment, including visual, auditory, and other sensor-based channels. For humans, this includes vision and proprioception; for machines, it includes cameras, LiDAR, radar, and GPS.
- **Attentional:** Mechanisms that determine which information is prioritized. In humans, this corresponds to cognitive focus; in AI systems, it maps to algorithmic attention mechanisms or sensor fusion strategies.
- **Learning:** Acquired knowledge or experience that influences decision-making. For humans, this includes prior driving experience and habitual behavior. For AI, it encompasses predictive models, skill policies, and learned representations.
- **Reactive:** Fast, often reflexive responses to stimuli, such as braking when an object suddenly appears.
- **Social:** Influences from external norms and policies, such as traffic rules, demographic trends, or culturally informed behaviors.
- **Environmental:** External contextual factors including road infrastructure, surrounding obstacles, pedestrian behavior, and weather conditions.

For modeling purposes, we distinguish between **internal** factors (e.g., sensory input, attention, learning, reactivity, social) and **external** disturbances (e.g., environmental conditions). Although the boundary is not always clearly defined (e.g., social norms may be both internalized and externally imposed), we use this distinction to guide the structure of our generative model. Each factor can be perturbed independently or jointly to simulate realistic anomaly scenarios. For instance, occluded sensors simulate perception failures; altered attention models distraction; and changes in environment (e.g., rain, congestion) induce external disturbances.

These factors serve as inputs to two key components governing the driver's behavior, each with a distinct role. In this framework, an **operator** is responsible for making high-level decisions, such as setting a reference state or target (cf. Table 2), based on policy and context. For instance, a driver may prefer different following distances depending on their experience or familiarity with a specific road. We define a *policy* as the mechanism that sets such reference states.

The **controller**, on the other hand, determines how to achieve the reference state based on real-time sensory inputs. It manages low-level actuation commands such as throttle, braking, and steering. The resulting control outputs, along with external environmental influences, cause changes in the system's state (e.g., a vehicle or pedestrian). These dynamics produce the *observable state*, which includes core kinematic and contextual variables.

### 5.1 Anomaly Generation Requirements

We expect simulations that supports both stochastic and rule-based generation of anomalies. Design objectives include:

(1) **Type diversity:** Coverage of different kinematic anomaly classes.
(2) **Realism:** Behavioral consistency before and after anomalies.
(3) **Controllability:** Fine-grained control over perturbation parameters.

Simulation outputs contain full observability of $Z$ and $X$ and are annotated with ground-truth anomaly labels. Detection models trained on these outputs rely only on observable features, supporting realistic deployment scenarios.

### 5.2 Generative Model Algebra

To formalize the systematic generation of kinematic anomalies, we introduce a concise algebraic framework that captures how various internal and external factors contribute to anomalous behavior in generative simulations. This abstraction allows anomaly configurations to be expressed as composable and modular elements.

Let:

- $o \in O$: a mobile object (e.g., pedestrian, vehicle, robot),
- $F = \{f_1, f_2, \ldots, f_n\}$: the set of generative factors influencing behavior (e.g., sensory, attentional, environmental),
- $C = \{c_1, c_2, \ldots, c_n\}$: the configuration space, where each $c_i$ corresponds to a specific instantiation or perturbation of factor $f_i \in F$,
- $\Gamma(o, F, C)$: a causally grounded anomaly generator producing a trajectory for object $o$ given factors $F$ and configurations $C$.

Each factor $f_i \in F$ represents a domain of influence over an agent's behavior, such as perception (e.g., sensory accuracy), cognitive state (e.g., attention, stress), or environment (e.g., weather, road conditions). The configuration $c_i \in C$ defines how that factor is instantiated or perturbed within the simulation.

**Definition 2** (Generative Process)**.** The generative process producing an anomalous trajectory is defined as:

$$X^*_{1:T} = \Gamma(o, F', C'),$$

where $F' \subseteq F$ is the subset of factors selected to induce anomaly, and $C'$ are their specific configurations (e.g., occluded vision, slippery road, distracted attention).

**Example 1** (Sensory Impairment in Human Driver)**.** Let

$$f^{\text{human}}_{\text{sensory}} \in F$$

denote the sensory factor for a human driver. A configuration modeling sensory impairment may be:

$$c_{\text{occlusion}} = \text{“reduced field of view”}.$$

An anomaly induced by visual occlusion is then represented as:

$$X^* = \Gamma\big(o, \{f^{\text{human}}_{\text{sensory}}\}, \{c_{\text{occlusion}}\}\big).$$

### 5.3 Algebraic Properties

Let $(\Gamma, \oplus)$ define an algebra over causal effects with a binary composition operator $\oplus$. An anomaly algebra should satisfy the following desirable properties:

- **Commutativity:** The order of composition of grounded anomalies does not matter, that is:

  $$\Gamma(o, F_1, C_1) \oplus \Gamma(o, F_2, C_2) = \Gamma(o, F_2, C_2) \oplus \Gamma(o, F_1, C_1).$$

  This property ensures that two anomalies resulting from the same causal effects, applied in different orders, yield the same outcome. For example, the anomaly created by applying causal factors $F_1 = f_{\text{attention}}$ with configuration $C_1 = \{\text{distracted}\}$, and $F_2 = f_{\text{environment}}$ with $C_2 = \{\text{wet road}\}$, should result in the same anomaly when applied to the same agent, regardless of order.
- **Associativity:** The grouping of causal effects does not affect the resulting anomaly:

  $$(\Gamma(o, F_1, C_1) \oplus \Gamma(o, F_2, C_2)) \oplus \Gamma(o, F_3, C_3)$$
  $$= \Gamma(o, F_1, C_1) \oplus (\Gamma(o, F_2, C_2) \oplus \Gamma(o, F_3, C_3)).$$

  This allows multi-factor anomalies to be constructed incrementally.
- **Idempotence:** Applying the same configuration to the same agent more than once should not alter the anomaly:

  $$\Gamma(o, F_1, C_1) \oplus \Gamma(o, F_1, C_1) = \Gamma(o, F_1, C_1).$$

  This property ensures that agents are not affected multiply by the same causal effect. For example, if an agent $o$ is already driving on a wet road using $\Gamma(o, f_{\text{environment}}, \text{wet road}\})$ caused by rain, then additional causal effect that causes the road to be wet due to a malfunctioning garden irrigation system using the same $\Gamma(o, f_{\text{environment}}, \text{wet road}\})$ should not affect the behavior of $o$.
- **Composability:** Anomalies induced by separate factor sets can be combined, e.g.,

  $$\Gamma(o, F_1, C_1) \oplus \Gamma(o, F_2, C_2) \implies \Gamma(o, F_1 \cup F_2, C_1 \cup C_2).$$

  This property allows multiple factors to be combined to create compound anomalies. For example,

  $$F' = \{f_{\text{attention}}, f_{\text{environment}}\}, \quad C' = \{\text{distracted}, \text{wet road}\}$$

  yields an anomaly resulting from both driver distraction and adverse road conditions. Note that the above relation is only one way due to the set union operator. For example, knowing that the compound anomaly $\Gamma_1 \oplus \Gamma_2$ causes a wet road does not imply whether $\Gamma_1$, $\Gamma_2$, or both individually cause the wet road condition.
- **Identity:** Let $e = A(o, \emptyset, \emptyset)$ represent the identity element (i.e., no causal effect), then:

  $$\Gamma(o, F, C) \oplus e = \Gamma(o, F, C).$$

  This property ensures that any observable effect is grounded in a causal factor. In other words, adding no causal effect should not change the resulting observable anomaly (or lack thereof).

This is a non-exhaustive list of algebraic properties that may (or may not be) desirable for an (envisioned) algebra over causal effects. But the main take-away from this section is that by defining causal effects as an algebra, we can defined operators to combine anomalies caused by one type of causal effects by anomalies caused by another type of causal effects to create complex compound anomalies that could be used in a simulation to create emerging anomalies that are not arbitrarily inserted into a simulation or dataset, but instead, emerge from causal factors and their composition.

*Semantic Typing.* Although the term 'semantic typing' is not always used explicitly, the underlying idea of assigning meaningful categories to influencing factors appears in several prior works. For instance, context-dependent anomaly detection frameworks utilize knowledge graph embeddings to encode the semantic context of factors and their relationships [48]. This approach enables richer reasoning about the origins and propagation of anomalies. Similarly, methods based on temporal, spatial, and semantic analysis of driving behavior treat behavioral deviations as transitions between semantically meaningful states [59]. This perspective supports both interpretability and structured reconstruction of anomalies.

Inspired by these approaches, we formalize semantic typing in our generative framework through a mapping

$$\tau(f_i) \in \{\text{driver-internal, vehicle-control, environmental}, \ldots\},$$

where $\tau(\cdot)$ is is a type function assigning each factor $f_i$ a semantic label that reflects its functional role within the mobility system. This categorization helps clarify how anomalies originate from different sources, such as human cognitive states, vehicle control dynamics, or environmental conditions. It also facilitates tracing how such anomalies influence various layers of behavior, from perception through planning to actuation.

This algebraic framework establishes a principled foundation for generating, categorizing, and experimenting with anomalies in simulation environments, supporting controlled benchmarking and analysis.

### 5.4 Examples of Generative Model Algebra

To illustrate the application of this algebraic framework in generating kinematic anomalies, we present representative examples aligned with the categories summarized in Table 3.

**Example 2** (Harsh Braking in Autonomous Vehicle)**.** Consider the acceleration factor:

$$f^{\text{vehicle}}_{\text{acceleration}} \in F,$$

and a configuration specifying an abrupt deceleration:

$$c_{\text{harsh}} = \text{“high negative acceleration spike”}.$$

The anomaly representing harsh braking is generated by:

$$X^* = \Gamma\big(o, \{f^{\text{vehicle}}_{\text{acceleration}}\}, \{c_{\text{harsh}}\}\big),$$

a core kinematic anomaly affecting acceleration.

**Example 3** (Illegal U-turn Combining Context and Kinematics)**.** Let

$$f_{\text{road}} \in F, \quad f^{\text{vehicle}}_{\text{heading}} \in F$$

represent road semantics and heading direction factors, respectively. The configurations can be:

$$c_{\text{no\_uturn}} = \text{“road segment disallowing U-turns”},$$
$$c_{\text{reverse\_heading}} = \text{“heading change of 180°”}.$$

The combined anomaly of performing an illegal U-turn is:

$$X^* = \Gamma\big(o, \{f_{\text{road}}, f^{\text{vehicle}}_{\text{heading}}\}, \{c_{\text{no\_uturn}}, c_{\text{reverse\_heading}}\}\big),$$

a compound anomaly involving both contextual and core kinematic variables.

**Example 4** (Unexpected Pedestrian Sprint)**.** Define

$$f^{\text{pedestrian}}_{\text{velocity}} \in F, \quad f_{\text{traffic\_signal}} \in F,$$

with configurations:

$$c_{\text{high\_speed}} = \text{“velocity exceeding normal pedestrian range”},$$
$$c_{\text{red\_light}} = \text{“crossing against red signal”}.$$

This anomaly models a pedestrian sprinting across an intersection during a red light:

$$X^* = \Gamma\big(o, \{f^{\text{pedestrian}}_{\text{velocity}}, f_{\text{traffic\_signal}}\}, \{c_{\text{high\_speed}}, c_{\text{red\_light}}\}\big).$$

Such compound anomalies capture complex behaviors combining kinematic and contextual deviations.

## 6 Conclusion

In this vision paper, we introduced the idea of a generative modeling framework for simulating kinematic anomalies in mobility contexts. By adopting a control-theoretic abstraction, we propose to map core behavioral components (including sensors, controllers, actuators, and disturbances) onto elements of vehicle and pedestrian dynamics. Building on this foundation, we propose an algebraic framework that formalizes how internal and external factors, such as attention, perception, and environmental conditions, can be composed to generate structured anomalous behaviors.

Unlike traditional anomaly detection approaches that treat anomalies as unstructured or rare outliers, our framework enables the deliberate synthesis of fine-grained, interpretable anomaly scenarios. This capability opens new directions for training robust autonomous systems, validating edge-case handling in simulators, and advancing driver assistance technologies through targeted behavioral perturbations.

While this paper focuses on the conceptual and formal underpinnings of the generative framework, we leave implementation and evaluation to future work. Our next step is to develop a simulation platform grounded in the proposed anomaly algebra. This platform will support parameterized anomaly injection across various modalities [23], including driving, cycling, and walking, enabling researchers to benchmark detection methods, assess control policies, and explore human-AI interaction in safety-critical settings.

One key challenge lies in modeling the composition of anomalies within simulations, particularly in understanding how different factors interact, whether additively, non-linearly, or in a context-dependent manner. For instance, how much lateral drift might result when a distracted driver encounters a wet road surface? While the proposed generative framework encapsulates the notion of composability, its practical implementation could be integrated with traffic simulation platforms such as SUMO [25] and CARLA [10], enabling realistic scenario generation and evaluation.

Although the primary focus is on generative modeling of kinematic anomalies, we plan to integrate this framework with agent-based patterns-of-life models [1, 24] to ensure behavioral grounding. Ultimately, we envision this framework as a testbed for principled, reproducible anomaly research across domains such as intelligent transportation, autonomous navigation, and human-centered mobility systems.

## References


[1] Hossein Amiri, Will Kohn, Shiyang Ruan, Joon-Seok Kim, Hamdi Kavak, Andrew Crooks, Dieter Pfoser, Carola Wenk, and Andreas Züfle. 2024. The Patterns of Life Human Mobility Simulation. In *Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems*. 653–656.

[2] Hossein Amiri, Ruochen Kong, and Andreas Züfle. 2024. Urban Anomalies: A Simulated Human Mobility Dataset with Injected Anomalies. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Geospatial Anomaly Detection*. 1–11.

[3] Zahra Amiri, Arash Heidari, Nima Jafari, and Mehdi Hosseinzadeh. 2024. Deep study on autonomous learning techniques for complex pattern recognition in interconnected information systems. *Computer Science Review* 54 (2024), 100666.

[4] John PK Bernstein, Nora Mattek, Katherine E Dorociak, Zachary T Beattie, Jeffrey A Kaye, John E Ferguson, and Adriana M Hughes. 2022. Age predicts older adults' driving self-regulation but not dangerous driving behaviors after controlling for executive function. *Gerontology* 68, 1 (2022), 98–105.

[5] Daniel Bogdoll, Maximilian Nitsche, and J Marius Zöllner. 2022. Anomaly detection in autonomous driving: A survey. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 4488–4499.

[6] Ann M Brewer. 2000. Road rage: what, who, when, where and how? *Transport reviews* 20, 1 (2000), 49–64.

[7] Basabi Chakraborty and Kotaro Nakano. 2016. Automatic detection of driver's awareness with cognitive task from driving behavior. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 003630–003633.

[8] Chen Chen, Xiaohua Zhao, Hao Liu, Guichao Ren, Yunlong Zhang, and Xiaoming Liu. 2019. Assessing the influence of adverse weather on traffic flow characteristics using a driving simulator and VISSIM. *Sustainability* 11, 3 (2019), 830.

[9] Shivoh Chirayil Nandakumar, Daniel Mitchell, Mustafa Suphi Erden, David Flynn, and Theodore Lim. 2024. Anomaly detection methods in autonomous robotic missions. *Sensors* 24, 4 (2024), 1330.

[10] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. 2017. CARLA: An Open Urban Driving Simulator. In *Proceedings of the 1st Annual Conference on Robot Learning*. 1–16.

[11] Ayman Elyoussoufi, Curtis L Walker, Alan W Black, and Gregory J DeGirolamo. 2023. The relationships between adverse weather, traffic mobility, and driver behavior. *Meteorology* 2, 4 (2023), 489–508.

[12] Pengcheng Fan, Jingqiu Guo, Yibing Wang, and Jasper S Wijnands. 2022. A hybrid deep learning approach for driver anomalous lane changing identification. *Accident Analysis & Prevention* 171 (2022), 106661.

[13] Mario Fiolić, Darko Babić, Dario Babić, and Sanja Tomasović. 2023. Effect of road markings and road signs quality on driving behaviour, driver's gaze patterns and driver's cognitive load at night-time. *Transportation research part F: traffic psychology and behaviour* 99 (2023), 306–318.

[14] Ralph Foorthuis. 2021. On the nature and types of anomalies: a review of deviations in data. *International journal of data science and analytics* 12, 4 (2021), 297–331.

[15] Qiang Gao, Xiaohan Wang, Chaoran Liu, Goce Trajcevski, Li Huang, and Fan Zhou. 2023. Open anomalous trajectory recognition via probabilistic metric learning. In *IJCAI*. International Joint Conferences on Artificial Intelligence Organization.

[16] Matias Garcia-Constantino, Paolo Missier, Phil Blythe, and Amy Weihong Guo. 2014. Measuring the impact of cognitive distractions on driving performance using time series analysis. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2047–2053.

[17] Beatriz González-Iglesias, José Antonio Gómez-Fraguela, and Mª Ángeles Luengo-Martín. 2012. Driving anger and traffic violations: Gender differences. *Transportation research part F: traffic psychology and behaviour* 15, 4 (2012), 404–412.

[18] Xiaolin Han, Reynold Cheng, Chenhao Ma, and Tobias Grubenmann. 2022. DeepTEA: Effective and efficient online time-dependent trajectory outlier detection. *Proceedings of the VLDB Endowment* 15, 7 (2022), 1493–1505.

[19] John D Hill and Linda Ng Boyle. 2007. Driver stress as influenced by driving maneuvers and roadway conditions. *Transportation Research Part F: Traffic Psychology and Behaviour* 10, 3 (2007), 177–186.

[20] Haoji Hu, Jina Kim, Jinwei Zhou, Sofia Kirsanova, JangHyeon Lee, and YaoYi Chiang. 2024. Context-Aware Trajectory Anomaly Detection. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Geospatial Anomaly Detection*. 12–15.

[21] Aparna Joshi, Kojo Adugyamfi, Jennifer Merickel, Pujitha Gunaratne, and Anuj Sharma. 2025. Driver Age and Its Effect on Key Driving Metrics: Insights from Dynamic Vehicle Data. *arXiv preprint arXiv:2501.06918* (2025).

[22] Lance Kennedy and Andreas Züfle. 2024. Kinematic Detection of Anomalies in Human Trajectory Data. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Geospatial Anomaly Detection*. 94–97.

[23] Joon-Seok Kim, Hamdi Kavak, Umar Manzoor, and Andreas Züfle. 2019. Advancing simulation experimentation capabilities with runtime interventions. In *2019 Spring Simulation Conference (SpringSim)*. IEEE, 1–11.

[24] Joon-Seok Kim, Gautam Malviya Thakur, Licia Amichi, Annetta Burger, Chathika Gunaratne, Joseph Tuccillo, Taylor Hauser, Joseph Bentley, Kevin Sparks, Debraj De, et al. 2024. Humonet: A framework for realistic modeling and simulation of human mobility network. In *2024 25th IEEE International Conference on Mobile Data Management (MDM)*. IEEE, 185–194.

[25] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, Laura Bieker, et al. 2012. Recent development and applications of SUMO-Simulation of Urban MObility. *International journal on advances in systems and measurements* 5, 3&4 (2012), 128–138.

[26] Prathyush Kumar Reddy Lebaku, Lu Gao, Jingran Sun, Xingju Wang, and Xuejian Kang. 2025. Assessing the Influence of Pavement Performance on Road Safety Through Crash Frequency and Severity Analysis: PKR Lebaku et al. *International Journal of Pavement Research and Technology* (2025), 1–22.

[27] Guofa Li, Yufei Yuan, Delin Ouyang, Long Zhang, Bangwei Yuan, Xiaoyu Chang, Zizheng Guo, and Gang Guo. 2023. Driver distraction from the EEG perspective: A review. *IEEE Sensors Journal* 24, 3 (2023), 2329–2349.

[28] Wenbin Li, Di Yao, Chang Gong, Xiaokai Chu, Quanliang Jing, Xiaolei Zhou, Yuxuan Zhang, Yunxia Fan, and Jingping Bi. 2024. Causaltad: Causal implicit generative model for debiased online trajectory anomaly detection. In *2024 IEEE 40th International Conference on Data Engineering (ICDE)*. IEEE, 4477–4490.

[29] Xishun Liao, Qinhua Jiang, Brian Yueshuai He, Yifan Liu, Chenchen Kuai, and Jiaqi Ma. 2024. Deep activity model: A generative approach for human mobility pattern synthesis. *arXiv preprint arXiv:2405.17468* (2024).

[30] Xishun Liao, Zhouqiao Zhao, Matthew J Barth, Amr Abdelraouf, Rohit Gupta, Kyungtae Han, Jiaqi Ma, and Guoyuan Wu. 2024. A review of personalization in driving behavior: Dataset, modeling, and validation. *IEEE Transactions on Intelligent Vehicles* (2024).

[31] Wei Liu, Min Hua, Zhiyun Deng, Zonglin Meng, Yanjun Huang, Chuan Hu, Shunhui Song, Letian Gao, Changsheng Liu, Bin Shuai, et al. 2023. A systematic survey of control techniques and applications in connected and automated vehicles. *IEEE Internet of Things Journal* 10, 24 (2023), 21892–21916.

[32] Yiding Liu, Kaiqi Zhao, Gao Cong, and Zhifeng Bao. 2020. Online anomalous trajectory detection with deep generative sequence modeling. In *2020 IEEE 36th International Conference on Data Engineering (ICDE)*. IEEE, 949–960.

[33] Felix Menze, Nathan W Churchill, Tom A Schweizer, and Simon J Graham. 2025. Simulated driving behavior over the adult age span. *Frontiers in Aging Neuroscience* 17 (2025), 1496224.

[34] Lisa J Molnar, David W Eby, JM Vivoda, SE Bogard, JS Zakraksek, RM St Louis, Nicole Zanier, Lindsay H Ryan, D LeBlanc, J Smith, et al. 2018. The effects of demographics, functioning, and perceptions on the relationship between self-reported and objective measures of driving exposure and patterns among older adults. *Transportation research part F: traffic psychology and behaviour* 54 (2018), 367–377.

[35] Kotaro Nakano and Basabi Chakraborty. 2019. Effective feature extraction from driving data for detection of danger awareness. In *2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST)*. IEEE, 1–6.

[36] Kotaro Nakano and Basabi Chakraborty. 2022. Real-time distraction detection from driving data based personal driving model using deep learning. *International Journal of Intelligent Transportation Systems Research* 20, 1 (2022), 238–251.

[37] Dimosthenis Pavlou, Gregory Christodoulou, and George Yannis. 2023. The impact of weather conditions and driver characteristics on road safety on rural roads. *Transportation research procedia* 72 (2023), 4081–4088.

[38] Iwona Pińskwar, Adam Choryński, and Dariusz Graczyk. 2024. Good weather for a ride (or not?): how weather conditions impact road accidents—a case study from Wielkopolska (Poland). *International journal of biometeorology* 68, 2 (2024), 317–331.

[39] Chittathuru Himala Praharsha and Alwin Poulose. 2024. CBAM VGG16: An efficient driver distraction classification using CBAM embedded VGG16 architecture. *Computers in biology and medicine* 180 (2024), 108945.

[40] Chinmay Vilas Samak, Tanmay Vilas Samak, and Sivanathan Kandhasamy. 2021. Control strategies for autonomous vehicles. In *Autonomous driving and advanced driver-assistance systems (ADAS)*. CRC Press, 37–86.

[41] Qiangqiang Shangguan, Ting Fu, and Shuo Liu. 2020. Investigating rear-end collision avoidance behavior under varied foggy weather conditions: A study using advanced driving simulator and survival analysis. *Accident Analysis & Prevention* 139 (2020), 105499.

[42] Wei Shao, Ziquan Fang, Lu Chen, and Yunjun Gao. 2025. Towards Trajectory Anomaly Detection: a Fine-Grained and Noise-Resilient Framework. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*. 2490–2501.

[43] Harpreet Singh and Ankit Kathuria. 2021. Analyzing driver behavior under naturalistic driving conditions: A review. *Accident Analysis & Prevention* 150 (2021), 105908.

[44] John Snow. 1855. *On the Mode of Communication of Cholera*. John Churchill.

[45] Imane Taourarti, Arunkumar Ramaswamy, Javier Ibanez-Guzman, Bruno Monsuez, and Adriana Tapus. 2025. Cross-Cultural Analysis of Car-Following Dynamics: A Comparative Study of Open-Source Trajectory Datasets. In *2025 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 330–337.

[46] Dimitrios I Tselentis and Eleonora Papadimitriou. 2023. Driver profile and driving pattern recognition for road safety assessment: Main challenges and future directions. *IEEE Open Journal of Intelligent Transportation Systems* 4 (2023), 83–100.

[47] Shiva Maleki Varnosfaderani, Mohd Rizwan Shaikh, and Mohamad Forouzanfar. 2025. A Comprehensive Review of Unobtrusive Biosensing in Intelligent Vehicles: Sensors, Algorithms, and Integration Challenges. *Bioengineering* 12, 6 (2025), 669.

[48] Nathan Vaska, Kevin Leahy, and Victoria Helus. 2022. Context-dependent anomaly detection with knowledge graph embedding models. In *2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)*. IEEE, 2020–2027.

[49] Boyu Wang, Wan Li, and Zulqarnain H Khattak. 2024. Anomaly detection in connected and autonomous vehicle trajectories using lstm autoencoder and gaussian mixture model. *Electronics* 13, 7 (2024), 1251.

[50] Chenhao Wang, Lisi Chen, Shuo Shang, Christian S Jensen, and Panos Kalnis. 2024. Multi-scale detection of anomalous spatio-temporal trajectories in evolving trajectory datasets. In *Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining*. 2980–2990.

[51] Lening Wang, Yilong Ren, Han Jiang, Pinlong Cai, Daocheng Fu, Tianqi Wang, Zhiyong Cui, Haiyang Yu, Xuesong Wang, Hanchu Zhou, et al. 2024. Accidentgpt: A v2x environmental perception multi-modal large model for accident analysis and prevention. In *2024 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 472–477.

[52] Natalie Watson-Brown, Verity Truelove, and Teresa Senserrick. 2024. Self-Regulating compliance to enhance safe driving behaviours. *Transportation research part F: traffic psychology and behaviour* 105 (2024), 437–453.

[53] Tianzheng Wei, Tong Zhu, Han Bai, Liang Zhao, and Xiuguang Wang. 2025. Effects of driver gender, driving experience, and visibility on car-following behavior. *Transportation Research Record* 2679, 1 (2025), 2166–2182.

[54] Hao Wu, Weiwei Sun, and Baihua Zheng. 2017. A fast trajectory outlier detection approach via driving behavior modeling. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. 837–846.

[55] Jian Xiao, Tianyuan Liu, and Genlin Ji. 2023. Human kinematics-inspired skeleton-based video anomaly detection. *arXiv preprint arXiv:2309.15662* (2023).

[56] Dongxiang Zhang, Zhihao Chang, Sai Wu, Ye Yuan, Kian-Lee Tan, and Gang Chen. 2020. Continuous trajectory similarity search for online outlier detection. *IEEE Transactions on Knowledge and Data Engineering* 34, 10 (2020), 4690–4704.

[57] Zheng Zhang, Hossein Amiri, Dazhou Yu, Yuntong Hu, Liang Zhao, and Andreas Züfle. 2024. Transferable Unsupervised Outlier Detection Framework for Human Semantic Trajectories. In *Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems*. 350–360.

[58] Zihao Zhang and Chenhui Liu. 2024. Exploration of the Heterogeneity among Elderly Drivers by Analyzing Traffic Crash Data: A Case Study in Pennsylvania, USA. *SAE International Journal of Transportation Safety* 12, 09-12-03-0008 (2024), 225–233.

[59] Rixing Zhu, Jianwu Fang, Hongke Xu, and Jianru Xue. 2019. Progressive temporal-spatial-semantic analysis of driving anomaly detection and recounting. *Sensors* 19, 23 (2019), 5098.

[60] Andreas Züfle, Dieter Pfoser, Carola Wenk, Andrew Crooks, Hamdi Kavak, Taylor Anderson, Joon-Seok Kim, Nathan Holt, and Andrew Diantonio. 2024. In Silico Human Mobility Data Science: Leveraging Massive Simulated Mobility Data (Vision Paper). *ACM Transactions on Spatial Algorithms and Systems* 10, 2 (2024), 1–27.