Trust on Trial: Navigating Accountability Diffusion in Autonomous Vehicle-Pedestrian Incidents Using Geels' Multilevel Perspective

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Harshal Nallapareddy

Fall 2024

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Kathryn A. Neeley, Associate Professor of STS, Department of Engineering and Society

I. Introduction

According to the Tesla Vehicle Safety Report, vehicles equipped with Tesla Autopilot technology are involved in nearly 90% fewer accidents than human-driven cars in the United States, highlighting the potential safety benefits of autonomous vehicles (Tesla, Inc., 2024, p. 2). As shown in Figure 1, In 2024, Tesla vehicles using Autopilot technology drive over seven million miles before experiencing an accident, compared to the national average of 750,000 miles. Moreover, as autonomous driving systems (ADSs) evolve to more advanced levels of autonomy, human driver intervention is considered a handicap, rather than support, to the safe operation of the AV, resulting in the ADS making its own driving decision on behalf of the human passengers. Despite the promising safety increase and crash prevention, there remains a significant uncertainty surrounding how accountability is taken in the event of an accident involving autonomous vehicles.

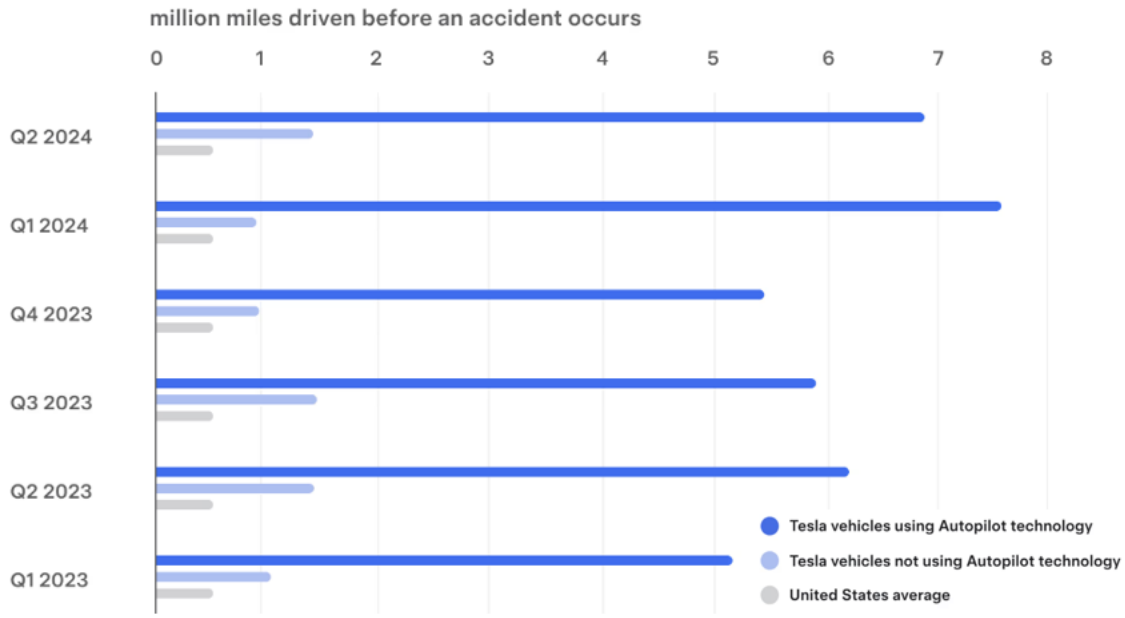

Figure 1: Tesla Vehicle Safety Report of number of miles in millions before crashing for Tesla vehicles with Autopilot, without Autopilot, and the United States average (Tesla Inc., 2024, p. 2).

If a fully autonomous vehicle causes an injury, death, or property damage, it is unclear whether the manufacturer, software designer, and/or driver should be held responsible. Given the diffusion of accountability, incidents involving autonomous vehicles become increasingly complex as these vehicles operate without a human driver and make decisions solely based on software, which ethically involves manufacturers and software designers who play a critical role in the vehicle's behavior (Himmelreich, 2018, pp. 1-4). Throughout this paper, I highlight the ethical intricacies in the sociotechnical transition from human-driven vehicles to autonomous vehicles, primarily focusing on accountability diffusion in accidents. Furthermore, I utilize Frank W. Geels' Multilevel Perspective to analyze the interplay between the niche, regime, and landscape levels of this transitions. In this paper, I argue that addressing accountability diffusion in autonomous vehicles accidents in crucial for public trust in the adoption of autonomous vehicles by overviewing the current literature regarding the ethics of autonomous driving technology.

II. Problem Definition: Accountability is blurred in Autonomous Vehicle and Pedestrian Incidents

Though autonomous vehicles are becoming more prevalent in urban environments, around 75% of American consumers remain cautious about riding in self-driving cars, which contributes to the negative public perception (Edmonds, 2019, p. 4). This apprehension is especially amplified in situations where the risk of harm is evident, such as traffic incidents that could affect passengers or other road users. Many drivers are particularly wary of relinquishing control to an autonomous system, fearing that the technology may not prioritize their safety as effectively as they would themselves (Shariff, Bonnefon, and Rahwan, 2017).

Pöllänen (2020, p. 1) found that when participants were asked to assign blame in vehicle accidents involving various actors—such as the driver, pedestrian, vehicle, manufacturer, and government—the outcome varied based on the level of autonomy. In accidents involving manual or semi-autonomous vehicles, drivers were primarily held responsible. However, in cases involving fully autonomous vehicles, the blame shifted significantly, with the vehicle manufacturers and the government being blamed more heavily, while the drivers themselves received less blame. This shift in accountability highlights a growing complexity in controversial ethical dilemmas. For instance, in situations where an autonomous vehicle must decide between protecting a passenger or a pedestrian, the vehicle is effectively making value judgments about human life. Such decisions further complicate the ethical landscape of accountability, as the autonomous system would be determining whose life to prioritize in critical moments (Gill et al., 2020, p. 1).

Specifically focusing on vehicle and pedestrian interactions, human drivers typically signal for pedestrians to either cross or wait at crosswalks. However, the current generation of autonomous vehicles lacks such explicit communication capabilities. To address this, researchers at Cornell University developed an Intent Communication System that provides visual messages to pedestrians, indicating whether it is safe to cross or wait (Matthews, et al., 2017, pp. 2-3). This advancement aims to bridge the communication gap between autonomous vehicles and pedestrians, ensuring safer interactions on the road.

In situations where individuals are fully in control of the vehicle, answering ethical questions about who to protect—passenger or pedestrian—can be extremely difficult. However, when riding in a fully autonomous vehicle, that sense of responsibility feels more removed. In a survey conducted by Gill (20202, pp. 1-3), respondents in such scenarios are more likely to

choose self-protection, even if it means putting the pedestrian in harm's way as depicted in Figure 2. This shift in accountability opens up possibilities for programming various ethical or moral frameworks into the vehicle's software to handle such dilemmas. For instance, from a utilitarian perspective, the vehicle might prioritize saving as many lives as possible, even if that puts the passenger's life at risk. Conversely, a virtue ethics framework could lead the vehicle to prioritize the safety of its passengers, deeming their lives more important than those outside the vehicle (McManus, Rutchick, 2019, p. 2). These differing frameworks reflect the ongoing debate about how autonomous systems should navigate complex moral decisions.

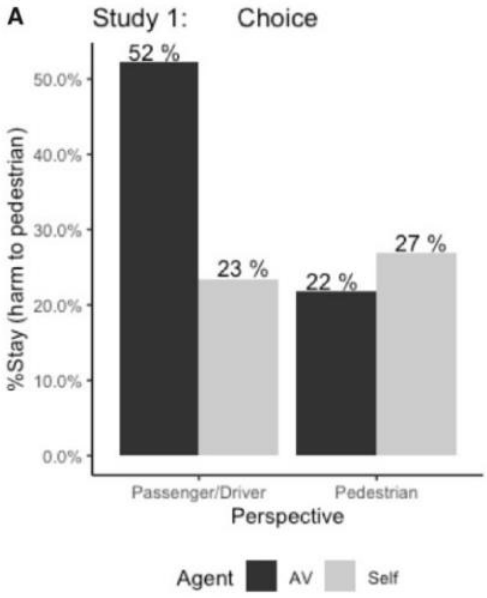

Figure 2: Survey responses indicating favorability towards self-protection in a vehicle over pedestrian protection. This shift in responsibility is making is easier for people to justify harm to a pedestrian when the vehicle is autonomous versus driven by a person (Gill, 2020, p. 3).

Despite the work accomplished in the ethics of autonomous vehicle technology, the knowledge gap of how accountability is assigned in the event of accidents involving these vehicles grows broad and deep. As the field continues to expand into uncharted territory,

understanding how accountability diffuses across various actors, such as manufacturers, software developers, and government regulators, is crucial. To address this gap, I turn to Frank W. Geels' multilevel perspective, which offers a framework for analyzing the sociotechnical transitions on multiple levels to examine how different societal levels interact to influence the evolving landscape of autonomous vehicles and the ethical challenges it presents.

III. Research Approach: Adopting Geels' Multilevel Perspective to Navigate Accountability Diffusion in Autonomous Vehicle Accidents

In "Transformations of Large Technical Systems", Geels introduces the multilevel perspective (MLP) of technological transitions, which examines how innovations like autonomous vehicles evolve through different layers of societal and technological influence. I chose Geels' framework because it offers a structured approach to understanding the layered complexity of autonomous vehicles in real-time transitions, providing valuable insights into how accountability diffusion can be navigated in AV-related accidents.

Describing Geels' Multilevel Perspective of Technological Transitions

Geels' work builds on existing STS research knowledge relating to large technical systems such as Hughes' concept of momentum being "the result of stabilizing connections between technology and society" (p. 124) and MacKenzie's understanding of technological trajectories as "self-fulfilling prophecies" (p. 124); however, he believes that there is "less attention to how momentum and stability are overcome" and further emphasizes that macro-level approaches abstract too much from case studies, and micro-level focus too much on the individual. To address this dichotomy, Geels proposes the "multilevel perspective on transitions from one system to another" (p. 126) that distinguishes three distinct levels because technological transitions are the outcome of "alignments between processes at these different

levels" (p. 126). Geels claims that understanding the dynamic interaction between the levels is vital for comprehending how large system transformations unfold and affect the "direction of *existing* trajectories" (p. 126).

Geels' multilevel perspective categorizes technological transition into three interacting levels known as the niche, regime, and landscape levels. The niche level refers to the experimental space where technological innovation is created and developed, the regime level represents the "ruleset or grammar" (p. 127) that governs the behavior of the sociotechnical system, and the landscape level is described as the "wider exogenous environment" (p. 129). This dynamic framework explains the complexities of surrounding accountability diffusion in autonomous vehicle accidents by examining the interplay to facilitate the smooth adoption of AVs. The evidence Geels incorporates primarily revolves around a historical case study of the Dutch highway system from 1950 to 2000, which he chronologically divides into critical periods that described the "technologies, rules, and interactions between the relevant social groups." The equivalent evidence relating to the research topic could include contemporary case studies of autonomous vehicle implementations and their regulatory challenges. Visualizing Geels' approach can be challenging due to the dynamic interplay of the three levels; however, it can be effectively represented in a diagram that illustrates the relationships and feedback loops among the niche, regime, and landscape levels, as shown below.

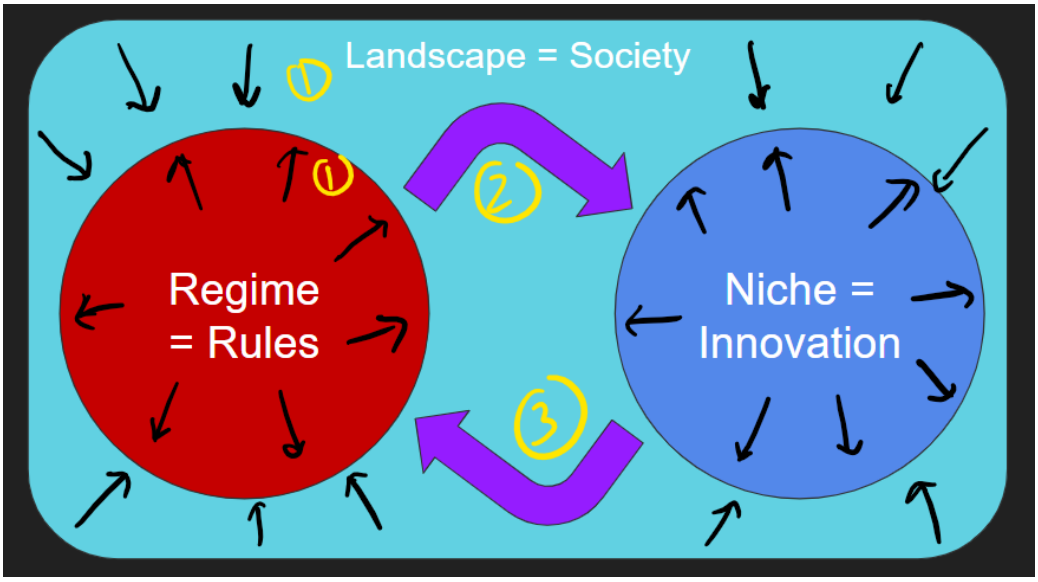

Figure 3: The interplay between the niche, regime, and landscape levels. The interplay between the levels can reveal relationships that aren't abundantly clear upon first look of a sociotechnical system (created by author).

The figure above illustrates the interplay between the niche, regime, and landscape levels, with arrows labeled (1) showing how societal and cultural pressures from the landscape level push for changes in the niche and regime levels, while these levels also push back. For example, AVs have the potential to revolutionize transportation, but public concerns about safety and liability have led to stricter regulations, limiting the scope of AV deployment. Arrows (2) and (3) depict the dynamic interaction between the niche and regime levels, where innovation reshapes established norms, and feedback from the regime level influences how innovation evolves. For instance, data from autonomous vehicle software may prompt regulators to modify traffic laws to accommodate AVs, while existing safety standards and liability laws can impose constraints that developers must adhere to, shaping the direction of technological advancements.

Case Study in Autonomous Vehicle and Pedestrian Accidents

As stated earlier, I narrowed the scope of the paper to autonomous vehicle and pedestrian accidents due to their unique challenges that raise ethical questions regarding autonomous

vehicle safety and decision-making in unpredictable environments. The specific case chosen was the October 2023 Cruise accident involving a pedestrian being trampled and dragged by a driverless vehicle created by a San Franciscan company named Cruise. This case, and the in-depth investigation report following the accident, holds significant relevance to present-day discussion on autonomous vehicle decision-making, regulatory policies, and public trust, and Geels' multilevel perspective can be employed to examine these layered dynamics to analyze how they can shape autonomous vehicle policy.

Key Steps in Analyzing the Evidence Using Geels' Multilevel Perspective

To analyze the accident, the investigation report is examined progressively through the three interconnected levels, niche, regime, and landscape, as shown in Figure 4. For the niche level, the focus is on identifying the technical failures that occurred within vehicle's detection and decision-making processes, which are critical actors in understanding the immediate causes of the accident. The Root Cause Analysis section of the report is reviewed to identify the gap between the detected environment and the resulting response in order to comprehend the vehicle's adverse behavior. Moving to the regime level, following Cruise's interactions with the Department of Motor Vehicles (DMV) and National Highway Traffic Safety Administration (NHTSA) will highlight the role of regulatory bodies and their enforced policies in shaping Cruise's future and could expose potential policy gaps. At the landscape level, the societal response to autonomous vehicles is inspected through anti-autonomous demonstrations, boycotts, and acts of vandalism in major cities. Analyzing how the three levels influence each other provides a holistic understanding of the complexities surrounding autonomous vehicle safety.

Figure 4: Flowchart describing the order of applying Geels' multilevel perspective to the case of the Cruise accident in October 2023 (created by author).

Appropriateness of Geels' Multilevel Perspective Framework

Geels' multilevel perspective has the promise of uncovering aspects of accountability diffusion that are not immediately obvious. By emphasizing the interactions between societal pressures, established norms, and emerging innovations, this framework sheds light on how regulations might redefine responsibilities among manufacturers, software designers, and drivers in the event of an accident. Moreover, it allows me to explore how public perceptions of autonomous vehicles can influence the technology's development and deployment trajectory. However, one limitation of Geels' perspective is its tendency to overlook individual behaviors and decisions, particularly critical in high-stakes scenarios like road accidents. Incorporating evidence from case studies, participant surveys, and interviews can capture the nuanced perspectives of various stakeholders and enhance the robustness of the analysis.

IV. Results: Gaps in Technical Capabilities, Regulation, and Public Trust Contribute to the Accountability Diffusion in Autonomous Vehicle Incidents

Cruise LLC, a subsidiary of General Motors, is a leader in autonomous vehicle technology and focuses on creating driverless electric vehicles to reduce urban traffic congestion and lower carbon emissions. Over the years, Cruise has attracted attention from the public and significant investors such as Honda, and it has conducted over five million miles of autonomous testing in major city streets such as San Francisco, Austin, and Phoenix while providing services

such as driverless taxis and goods delivery. In October 2023, a tragic incident in San Francisco occurred involving one of Cruise's autonomous vehicles. A woman crossing the street was struck by another car and then, while lying on the road, was run over and dragged by a Cruise vehicle. The Cruise vehicle reportedly braked and attempted to stop after detecting the woman but only after dragging her 20 feet. This incident sparked widespread debate on the safety protocols and accountability associated with fully autonomous vehicles and led to increased scrutiny and public backlash of Cruise's technology and the autonomous vehicles industry.

In the following months, Quinn Emmanuel Trial Lawyers investigated on the accident and reported on the technical failures leading up to the accident, Cruise's communication with regulatory authorities, and the societal pressure for Cruise's license suspension. At the niche level, the technical breakdown conducted by Exponent Inc., an engineering consulting firm that conducted a third-party root cause analysis of the accident, revealed that the vehicle's sensors and decision-making algorithms failed to correctly identify and assess the pedestrian's presence and trajectory, resulting in a delayed response that exacerbated the situation. An analysis of the action-to-action in Figure 3 illustrates how the vehicle processed information leading up to and during the collision and highlights key moments where detection and decision delays likely contributed to the collision.

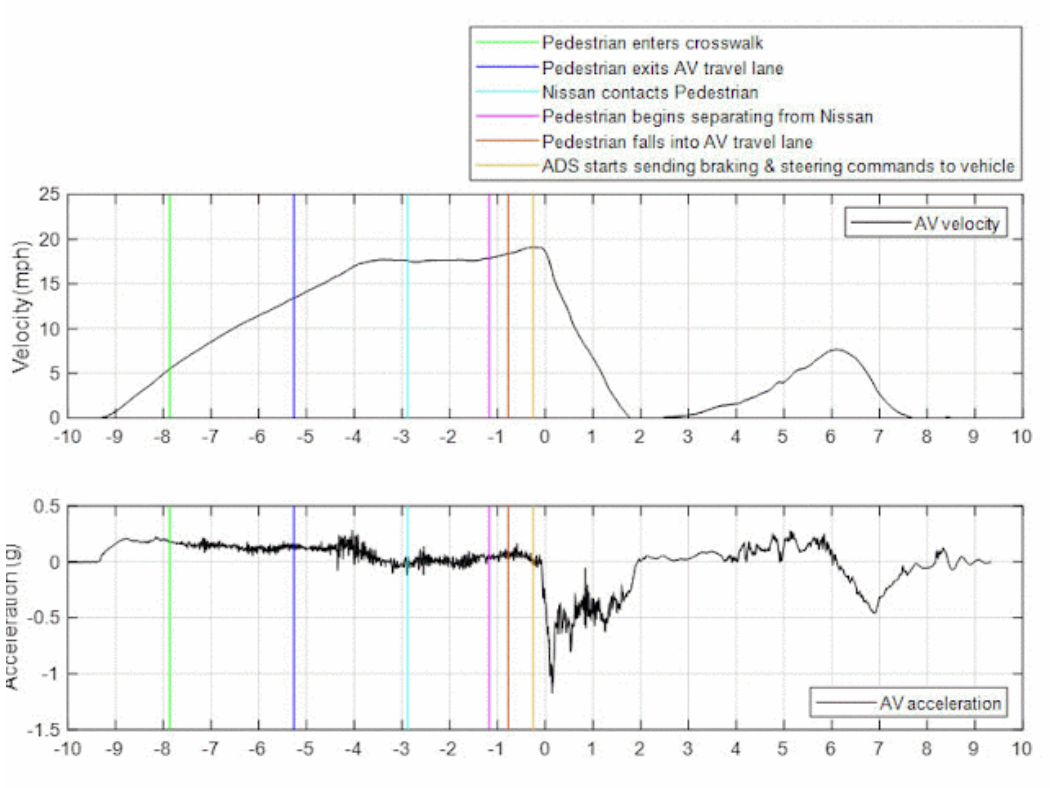

Figure 5: Action-to-action graph of key events where the x-axis is time before and after the contact point between the autonomous vehicle and pedestrian (t=0s), and the y-axis is the velocity/acceleration of the Cruise vehicle. The events are color coordinated and follow in order as presented in the legend (Quinn Emmanuel, 2024).

The analysis suggests that while Cruise's system is optimized for general road navigation, limitations persist in handling unpredictable pedestrian behavior, especially in complex urban environments. This gap demonstrates that while technical advances have been made, autonomous vehicles still require significant improvement at the niche level to safely and efficiently manage pedestrian interactions.

In response to the accident, Cruise was required to report to several regulatory bodies such as the DMV and NHTSA. Cruise's series of disclosures exposed discrepancies between autonomous vehicle company practices and regulatory expectations. For instance, the DMV's suspension order on October 24th suggests that the current regulatory landscape lacks sufficient safeguards or standards that account for autonomous vehicle accidents involving pedestrians. In

a meeting with the San Francisco Municipal Transportation Agency (SFMTA), the video of the accident after the collision was presented in which the “very obvious” pullover maneuver was attempted to stabilize the vehicle in a “minimal risk condition”, a DMV requirement for all autonomous vehicles to come to a safe controlled stop, to which a government official responded, “so the vehicle comes to a stop and continues driving?” This discourse between Cruise and the SFMTA reflects the lack of cohesiveness between practice and expectation, which broadens the significant oversights that affect autonomous vehicle deployment in real-world urban settings. This incident underlines how regulatory frameworks need to adapt to the unique challenges posed by fully autonomous vehicles.

At the landscape level, societal perceptions and public trust in autonomous vehicles play a crucial role in shaping regulatory and commercial responses. The immediate public reaction to the accident underscored a strong societal demand for autonomous vehicle accountability and safety transparency. Media coverage quickly amplified the story, focusing on the vehicle’s failure to respond adequately in a pedestrian scenario as reported in ABC: vandalism of a Waymo vehicle in San Francisco’s Chinatown, where an organized group, frustrated with the presence of autonomous vehicles, set a vehicle on fire during the Lunar New Year celebrations. This wave of negative attention not only pressured Cruise to publicly address safety concerns but also influenced regulatory actions, which ultimately led to the suspension of Cruise’s operating license by Californian authorities. The public’s strong reaction demonstrates an ongoing societal hesitancy toward autonomous vehicles that could affect their widespread acceptance if companies and regulators fail to address these concerns adequately.

Level	Key Findings	Evidence
Niche	Technical limitations in the vehicle's detection and decision-making systems contributed to the delayed response in pedestrian detection and braking.	Exponent Inc. root cause analysis and action-to-action graph showing detection and decision delays in the vehicle's response sequence.
Regime	Regulatory gaps and misalignment between regulatory expectations and AV operational capabilities led to scrutiny of Cruise's practices post-incident.	DMV and NHTSA interactions with Cruise; comments from SFMTA questioning the vehicle's failure to execute a controlled stop as mandated by DMV.
Landscape	Societal backlash and distrust toward autonomous vehicles, especially after high-profile incidents, intensify pressure on AV companies to ensure safety.	Media coverage of anti-AV sentiment, including incidents of vandalism (e.g., Waymo vehicle set on fire), reflecting public frustration and impacting regulatory actions.

Table 1: Key findings and associated evidence for each level as described by Geels' multilevel perspective. Explicitly identifying each of the levels is important for properly examining the interplay between the levels. Created by Author.

Explicitly identifying the levels of Geels' multilevel perspective allows for the examination of the interplay between the levels as previously described in Figure 3. The niche-level technical limitations – detection and decision-making failures – that resulted in the devasting accident can directly impact how regulatory bodies, like the DMV and NHTSA, scrutinize the current state and future progress of autonomous driving systems and could adjust enforceable policy at the regime level by establishing clear, actionable standards that shape operational protocols and safety requirements for future autonomous vehicle development. In turn, breakthroughs in autonomous driving research can place pressure on regulatory bodies to adapt and update existing frameworks while certain leaps may even redefine how safety and liability are practically considered. Likewise, public reactions captured at the landscape level such as negative media attention and violent activist groups can intensify regulatory pressure and created societal demand for accountability on the corporate level, driving stricter oversight and comprehensive safety features. Together, these levels reveal a feedback loop where technical

failures/breakthroughs illicit regulatory responses, which further influence and are influenced by societal pressures, and societal demand necessitates further technological advancement.

V. Conclusion

In this paper, I argue that addressing accountability diffusion in autonomous vehicle accidents is crucial for building and maintaining public trust as autonomous vehicles move toward widespread adoption. Through an in-depth analysis of the October 2nd Cruise incident, the study highlights how the lack of clear accountability—spanning technical, regulatory, and social levels—reveals complex, unresolved challenges in autonomous vehicle governance. By using Geels' multilevel perspective, this research underscores the need for transparency and regulatory clarity to bridge gaps between AV developers, policymakers, and public perception. A potential practical application for the research could be the development of clear, enforceable standards – such as an autonomous vehicle legal code – to strengthen regulatory oversight and provide a rulebook for autonomous vehicles developers. Such standards could also encourage broader social acceptance by providing the public with assurance of reliable course of action in case of autonomous vehicle accidents. However, a key limitation of this approach is the focus on autonomous vehicle versus pedestrian incidents. In other cases, such as autonomous vehicle-to-vehicle incidents, adds a layer of complexity that further diffuses accountability and brings difficulty in understanding failure modes. Expanding future research to include autonomous vehicle-to-vehicle and account for inter-vehicle communication and shared environment data could yield more comprehensive insights into accountability frameworks for autonomous vehicles.

References:

Emanuel, Q. (2024, January 24). Report to the boards of directors of Cruise LLC, GM CRUISE Holdings LLC, AND General Motors Holdings LLC regarding the October 2, 2023 accidents in San Francisco. Ctfassets.net. https://assets.ctfassets.net/95kuvdv8zn1v/1mb55pLYkkXVn0nXxEXz7w/9fb0e4938a89dc5cc09bf39e86ce5b9c/2024.01.24_Quinn_Emanuel_Report_re_Cruise.pdf [report]

Geels, F. (2007). Transformations of large technical systems: A multilevel analysis of the Dutch highway system (1959-2000). *Science, Technology, and Human Values* 32(2), 123-149.

Gill, T. (2020). Blame it on the self-driving car: how autonomous vehicles can alter consumer morality. *Journal of Consumer Research*, 47(2), 272-291.

Himmelreich, J. (2018). Never mind the trolley: The ethics of autonomous vehicles in mundane situations. *Ethical Theory and Moral Practice*, 21(3), 669-684.

Hubbard, S. M. L. (2018). Automated vehicle legislative issues. *Transportation Research Record*, 2672(7), 1-13. https://doi.org/10.1177/0361198118774155.

Koopman, P. (2024, January 29). The Exponent Report on the Cruise Pedestrian Dragging Mishap. Blogspot.com. https://safeautonomy.blogspot.com/2024/01/the-exponent-report-on-cruise.html. [blog]

Liu, L., Lu, S., Zhong, R., Wu, B., Yao, Y., Zhang, Q., & Shi, W. (2020). Computing systems for autonomous driving: State of the art and challenges. *IEEE Internet of Things Journal*, 8(8), 6469-6486.

Matthews, M., Chowdhary, G., & Kieson, E. (2017, August 23). Intent Communication between Autonomous Vehicles and Pedestrians. *ArXiv*. https://doi.org/10.48550/arXiv.1708.07123

McManus, R. M., & Rutchick, A. M. (2019). Autonomous vehicles and the attribution of moral responsibility. *Social Psychological and Personality Science*, 10(3), 345-352. https://doi.org/10.1177/1948550618755875.

Pöllänen, E., Read, G. J. M., Lane, B. R., Thompson, J., & Salmon, P. M. (2020). Who is to blame for crashes involving autonomous vehicles? Exploring blame attribution across the road transport system. *Ergonomics*, 63(5), 525–537. https://doi.org/10.1080/00140139.2020.1744064.

Tesla Inc. (2024, October). Tesla vehicle safety report. *Tesla.com*. https://www.tesla.com/VehicleSafetyReport.

Waymo car set on fire, vandalized in San Francisco's Chinatown, video shows. (2024, February 12). *ABC7 Los Angeles*. https://abc7.com/waymo-car-set-on-fire-san-francisco-sf/14415284/. [news article]