Written Testimony of Dr. Henry Liu

Professor of Civil and Environmental Engineering, University of Michigan
Director, University of Michigan Transportation Research Institute (UMTRI)
Director, Center for Connected and Automated Transportation (CCAT)
Director, Mcity

Prepared for

U.S. House of Representatives
House Committee on Science, Space, and Technology
Subcommittee on Research and Technology

Hearing:
"Accelerating Progress: U.S. Surface Transportation Research"

Wednesday, February 11, 2026, 2:00 p.m.
2318 Rayburn House Office Building

Good afternoon, Chairman Obernolte, Ranking Member Stevens, and Members of the Subcommittee:

My name is Henry Liu. I am a Professor of Civil and Environmental Engineering at the University of Michigan, and I serve as Director of the University of Michigan Transportation Research Institute (UMTRI), Mcity, and the Center for Connected and Automated Transportation (CCAT). Thank you for the opportunity to appear before you to discuss how the United States can accelerate progress in surface transportation research. UMTRI is one of the nation's leading academic transportation research institutes, dedicated to advancing safety and mobility through interdisciplinary research and public-private partnership. The Mcity Test Facility was the world's first purpose-built proving ground for connected and automated vehicles, and through the NSF-funded Mcity 2.0 effort, it has expanded into a remote-accessible, mixed-reality testbed for autonomous vehicle (AV) evaluation. CCAT is the USDOT-funded Region 5 University Transportation Center and the designated regional center for connected and automated vehicles and infrastructure, with a statutory focus on safety.

I appreciate the Subcommittee's focus on strengthening the nation's surface transportation research enterprise, especially at a moment when the stakes could not be higher: roadway deaths and serious injuries remain unacceptably high; supply chains and travel reliability are strained; and artificial intelligence is rapidly transforming the technologies that shape how people and goods move. The United States can lead in transportation innovation, but leadership requires more than basic research and technology invention; it also requires validation, deployment, and measurable public benefit. In other words, we must strengthen the research-to-impact pipeline.

My testimony focuses on two priority areas for strengthening U.S. surface transportation research and translating it into real-world safety and mobility impact:

1. **Research to establish a rigorous national framework for AV safety testing and evaluation.**

   The United States needs an AV safety evaluation framework that is scientifically grounded, comparable across systems, and credible to the public. The core challenge is not just "can the vehicle drive safely?" but how to generate trustworthy safety evidence that support broad deployment—evidence that addresses rare, high-severity crash scenarios and helps build consumer confidence.

2. **Research to advance digital road infrastructure to improve safety and mobility with "do more with less" capabilities.**

   The fastest, most scalable near-term impact often comes from modernizing the infrastructure we already have—especially intersections and corridors—using data, sensors, communications, and AI-enabled operations. The goal is to enable agencies to do more with less and achieve measurable improvements without requiring extensive and expensive new physical infrastructure.

Why now? AI is transforming many sectors of society, but transportation is uniquely consequential because it is safety-critical, ubiquitous, and tightly coupled to physical reality. As AI capabilities advance, transportation can benefit from dramatic improvements in safety and efficiency, but only if we build the testing, evaluation, and infrastructure foundations that allow innovation to scale responsibly.

In practical terms, we are approaching a threshold moment in both priority areas:

- AV capabilities are improving rapidly and deployments are expanding; yet public concern remains high, in part because independent, comparable safety evidence is still limited.

- For infrastructure, agencies face fiscal and workforce constraints; data-driven operations can stretch limited resources while improving safety outcomes.

Today, I will describe the importance of each priority area, the cross-cutting barriers Congress can address, and my recommended actions.

1. Priority Area 1: AV evaluation for scaled deployment

Why is a rigorous AV evaluation framework needed

As autonomous driving systems move from limited pilots toward broader deployment, society needs a credible way to answer a basic question: how do we know a system is safe enough to scale? Without a rigorous, unbiased evaluation framework, we face two unacceptable outcomes, either premature deployment based on incomplete evidence that can harm the public and undermine trust, or delayed adoption of potentially life-saving technology because regulators, agencies, insurers, and communities lack a principled basis for confidence. Credible evidence that AVs are safe is also in the public interest. Consumers cannot independently verify safety claims, and local and federal decision-makers need transparent, auditable evidence to govern operations responsibly and consistently.

Why it is difficult

AV safety evaluation is difficult not only because crashes are rare and the operational world is effectively open-ended (known as "Curse of Rarity"[1]), but also because modern AV stacks are increasingly AI-driven and partially black-box, making their internal reasoning hard to inspect, reproduce, or certify with traditional engineering methods. Performance can vary sharply with small changes in context, and learning-based components can fail in non-intuitive ways, including producing confident but incorrect outputs (often described as "hallucinations" in AI systems), especially under distribution shift, sensor degradation, rare corner cases, or adversarial interactions. Safety is also highly dependent on the specific conditions under which an AV will

---

[1] Liu, H. and Feng, S. (2024) Curse of rarity for autonomous vehicles, Nature Communications 15, 4808. https://www.nature.com/articles/s41467-024-49194-0

operate – what is known as its operational design domain, or ODD. A system that performs well on a limited set of routes, speeds, or weather conditions may not be safe when expanded, and software updates and model drift mean that safety is not a one-time assessment problem but an ongoing challenge of versioning and continuous monitoring, making "just drive more miles" an impractical and ethically insufficient validation strategy.

Why extensive research is essential

Because of these fundamentals, a national AV evaluation framework cannot be created by policy alone; it requires sustained research to make safety evidence rigorous, scalable, and comparable. We need new methods to quantify safety under rarity and uncertainty; scenario-based approaches that define meaningful coverage of real-world risk; credible connections between simulation, closed-course testing, and limited on-road exposure; and standardized metrics and benchmarking protocols that enable independent assessment without forcing unnecessary disclosure of proprietary details. In short, extensive research is what turns AV evaluation from ad hoc reporting into a scientific, repeatable process that can support both innovation and public trust on a national scale.

Mcity's AV Safety Assessment Program

Mcity's AV Safety Assessment Program[2] offers a practical blueprint for third-party, performance-based assessment centered on behavioral and probabilistic safety. It comprises two complementary components: First, Basic Behavioral Competency Testing provides an ODD-specific scenario library to verify foundational driving behaviors, both routine and safety-critical, using risk-tiered scenarios and standardized metrics, establishing a behavioral "floor" analogous to human licensing prior to large-scale deployment. Second, the Driving Intelligence Test provides a statistical evaluation of overall safety in a naturalistic and adversarial environment that uncovers previously unknown unsafe behaviors and estimates risk as a rate with confidence bounds (e.g., crashes or hazardous conflicts per mile) for a defined ODD, accelerating exposure to rare events without bias. Together, these components address both the "known" competencies and the long-tail risks that matter most to the public, while producing unbiased outputs that support transparent reporting and can be integrated into a safety case.

The program is still a developing framework, in that its scenario libraries, benchmarking thresholds, and evidence-integration protocols require further research and multi-stakeholder validation to mature into a safety evaluation approach that can be broadly adopted.

2. **Priority Area 2: AI for Digital Infrastructure: why it is needed and where the opportunities are**

---

[2] Liu, et. al. (2025) Behavioral Safety Assessment towards Large-scale Deployment of Autonomous Vehicles, https://arxiv.org/abs/2505.16214.

Digital infrastructure refers to the software-defined, data-enabled layer of the roadway system, including connected sensors, communications (e.g., C-V2X), modern signal controllers, edge/cloud computing, and standardized data interfaces, that allow agencies to measure conditions in real time and operate roads more safely and efficiently. AI-enabled digital infrastructure is needed because most safety and mobility gains in the near term will come not from rebuilding roads, but from making existing roads smarter, especially intersections and corridors where risk, delay, and emissions concentrate. State and local agencies are asked to do more with limited budgets and staffing, yet they operate systems that already generate valuable data (signals, detectors, cameras, connected-vehicle messages, maintenance logs). The opportunity is to turn that fragmented data into actionable, real-time operational intelligence: detecting risk earlier, optimizing control strategies continuously rather than episodically, prioritizing maintenance and upgrades based on measurable safety benefit, and creating a digital backbone that supports both human-driven and automated vehicles. In other words, digital infrastructure is the fastest path to scalable safety impact because it can be deployed widely, updated quickly, and evaluated objectively.

Example 1: Traffic signal optimization delivers measurable benefits without new equipment

Traffic signals are one of the highest-leverage control points in the entire surface transportation system. With AI and data-driven operations, agencies can optimize timing plans more frequently, adapt to recurring congestion patterns, and reduce unnecessary stops, often without major new roadside equipment. Funded by CCAT, Road Commission of Oakland County (RCOC), and Michigan Department of Transportation (MDOT), UMTRI has developed a new signal timing optimization system[3] by leveraging anonymized vehicle trajectory data. The system was tested in Birmingham, Michigan in 2023, and reduced delays by up to 20% and stops by up to 30%. More recently, we prototyped this system through a USDOT Stage 1 SMART Grant awarded to RCOC. We computed signal performance measures for over 1,400 signalized intersections across Oakland County and identified 40 intersections across seven corridors for retiming. Early results from two corridors show approximately 20% reductions in both delay and number of stops across all movements, with some time-of-day periods achieving reductions of up to 30% in delay and 40% in stops. This "**do more with less**" approach can scale quickly across jurisdictions and deliver near-term improvements while building the data foundation for future connected and automated operations.

Example 2: Near-miss detection proactively monitors safety before crashes happen

Crashes are too rare to serve as the sole feedback signal for safety management, especially at a single intersection or corridor. AI creates an opportunity to measure safety proactively using near-misses and hazardous interactions that occur far more frequently than crashes. Funded by

---

[3] Wang, et al. (2024) Traffic light optimization with low penetration rate vehicle trajectory data, *Nature Communications* 15, 1306. https://www.nature.com/articles/s41467-024-45427-4

Mcity and the USDOT Smart Intersections Project, UMTRI has developed a video-based roadside perception system, MSight[4], to extract multimodal trajectories from existing cameras and compute standardized conflict indicators. Agencies can detect rising risk, identify which movements and behaviors are driving it, and evaluate whether countermeasures actually reduce dangerous conflicts, weeks or months before crash statistics would show a change. This turns safety from a retrospective process into a continuous risk-monitoring function, enabling faster, evidence-based interventions. This system is currently being deployed at more than 50 intersections in the City of Ann Arbor, Michigan.

3. Barriers to AI for Transportation research

The topic of this hearing title includes the word "accelerating," and that is exactly right. The largest opportunities to accelerate progress are often not in generating new ideas, but in strengthening the systems that convert those ideas into measurable safety and modernization outcomes.

Based on my experience and consistent with my outline, three barriers deserve particular attention.

Barrier 1: No dedicated federal program to support AV testing and evaluation research

Despite rapid progress in automated driving and growing public demand for trustworthy safety evidence, the United States still lacks a dedicated, sustained federal research program focused specifically on AV testing, evaluation, and safety assurance, i.e., the science of how to measure safety, quantify uncertainty, validate simulation and testbeds, and establish repeatable benchmarking protocols across operational design domains. As a result, AV safety evaluation research is funded in a fragmented way, often as small components of broader programs, making it difficult to build shared scenario libraries, common metrics, data standards, and independent assessment capability at national scale. A dedicated program would not regulate technology; rather, it would strengthen the evidence infrastructure needed for transparent, comparable, and auditable evaluation, so innovations can scale responsibly and public trust can be earned through measurable safety performance.

Barrier 2: Academic institutions are talent-rich, but constrained in data and compute

Transportation research now depends on large-scale data, advanced models, and compute-intensive testing. Yet universities are increasingly outmatched by the scale of proprietary datasets and high-performance compute available in the private sector, creating a structural gap that limits independent evaluation, reproducibility, and the pace of scientific progress. To close this gap, the

---

[4] Zhang, et al. (2023) MSight: An Edge-Cloud Infrastructure-based Perception System for Connected Automated Vehicles, https://arxiv.org/abs/2310.05290

nation should build shared, research-accessible infrastructure: a central, privacy-preserving transportation data repository (with standardized formats, governance, and secure access) and dedicated compute facilities (e.g., GPU-enabled resources and cloud/edge credits) that qualified researchers can use for model development, scenario-based testing, and rigorous validation.

This matters because universities provide three public-interest functions that cannot be left to commercial incentives alone: independent evaluation of safety claims (the "voice of reason" role); workforce development, training engineers and scientists fluent in both transportation systems and AI; and open science and reproducibility, which become more important as systems grow more data-intensive and opaque.

Barrier 3: Lack of stable support for testbeds and living labs

Transportation is not purely digital. For AI-enabled transportation, especially autonomy and infrastructure sensing, systems must be tested in environments that reflect physical reality, failure modes, weather, human behavior, and operational constraints.

When testbeds are treated as short-term projects rather than sustained national assets, we get predictable failures: promising research that cannot be validated; fragmented pilots with non-comparable results; and overreliance on simulation without sufficient grounding.

4. Funding and program recommendations

After prioritizing the "what" and the "why," let me close with the "how."

Establish federal programs to support AV testing and evaluation research

Establishing dedicated federal programs to support AV testing and evaluation research is essential to build the evidence infrastructure needed for scaled AV deployment. A targeted federal program, coordinated across USDOT, NSF, and other relevant agencies, should fund the science of evaluation itself: rare-event risk estimation; uncertainty quantification; surrogate safety measures; validation of simulation and digital twins against real-world outcomes; and repeatable protocols that integrate evidence from simulation, closed-course testing, and limited on-road pilots for defined operational design domains. The goal is not to pick technology winners, but to enable transparent, comparable, auditable safety evidence, so innovation can scale responsibly and public trust can be earned through measurable performance.

Build shared, research-accessible infrastructure for data and compute

To enable rigorous AI-for-transportation research at national scale, the United States should build shared, research-accessible infrastructure that reduces the structural gap between academic researchers and the resources available in the private sector. This should include a central, privacy-preserving transportation data repository that curates high-value datasets (e.g., roadway operations, trajectories, safety events, connected infrastructure messages) under standardized

formats, clear governance, and secure access, so researchers can reproduce results, benchmark methods, and develop evaluation protocols without compromising privacy or security. In parallel, the nation should provide dedicated compute facilities, including GPU-enabled resources and cloud/edge credits, so qualified researchers can train and test large models, run scenario-based simulation at scale, and perform compute-intensive validation and uncertainty analysis. Together, these shared data-and-compute resources would function as "evidence infrastructure," accelerating research-to-impact while strengthening transparency, independence, and public trust.

Provide stable support for testbeds and living labs

Providing stable, long-term support for testbeds and living labs is essential to close the gap between promising research and verified real-world impact. Transportation is a safety-critical cyber-physical system, so evaluation cannot rely on simulation alone or on short-term pilots that are difficult to reproduce. It requires sustained environments where new technologies can be tested under controlled, repeatable conditions and then validated in realistic operational settings. A stable federal commitment, covering both operations and modernization, would allow testbeds and living labs to maintain instrumentation, data pipelines, cybersecurity protections, and safety protocols; to develop shared scenario libraries and evaluation standards; and to provide equitable access for researchers, agencies, and smaller jurisdictions that cannot build these capabilities independently. In practice, sustained support turns testbeds into national assets that enable independent assessment, accelerate translation, and ensure innovations are deployed responsibly with evidence the public can trust.

Reauthorize the University Transportation Centers (UTC) Program

Reauthorizing the University Transportation Centers (UTC) Program is one of the most consequential steps Congress can take to strengthen U.S. surface transportation research and ensure that innovation translates into measurable public benefit. The UTC program is the nation's backbone for the research-to-workforce-to-deployment pipeline. It supports high-impact, multidisciplinary research; trains the next generation of transportation professionals; and builds durable partnerships among universities, state and local agencies, and industry so that new methods can move from publications into practice. In an era when transportation is rapidly becoming an AI-enabled cyber-physical system, this pipeline is more important than ever, because safety, trust, and operational performance increasingly depend on rigorous evaluation, high-quality data, and expertise that spans engineering, computing, and public policy. Reauthorization is therefore not simply a continuation of existing funding; it is a strategic national commitment to sustain and modernize the UTC ecosystem so the United States can lead in transportation safety, infrastructure modernization, and the responsible deployment of emerging technologies at scale.

5. Closing

In closing, the United States can accelerate progress in surface transportation research by focusing national leadership on two outcomes: a rigorous, evidence-based framework for AV safety testing and evaluation that earns public trust through independent, unbiased performance evidence; and AI-enabled digital infrastructure modernization that delivers measurable safety and mobility improvements at scale. Achieving these outcomes requires strengthening the research-to-impact pipeline, the part of the system that turns promising ideas into validated, deployable solutions.

To do that, I respectfully urge Congress to: establish dedicated federal programs to advance the science of AV evaluation; build shared, research-accessible infrastructure for transportation data and compute; provide stable, long-term support for testbeds and living labs as national assets; and reauthorize and modernize the University Transportation Centers Program as the backbone of the research-to-workforce-to-deployment pipeline. With these steps, the United States can lead in transportation safety, infrastructure modernization, and the responsible deployment of AI-enabled mobility technologies, grounded in rigorous evidence and measurable public benefit.

Thank you. I look forward to your questions.

Appendix: Introduction to UMTRI/CCAT/Mcity

The **University of Michigan Transportation Research Institute (UMTRI)** was established in 1965, and is one of the world's leading academic transportation research institutes, dedicated to advancing safety and mobility through interdisciplinary research and public-private partnership. UMTRI's work spans crash and injury prevention, driver behavior and human factors, connected and automated vehicle safety, and large-scale mobility analytics, combining scientific rigor with partnerships that move research into practice. UMTRI built and maintains the Ann Arbor Connected Environment—funded through a series of USDOT projects—which deploys C-V2X roadside units and perception sensors at approximately 75 intersections, paired with an edge-cloud analytics platform capable of near-miss detection. This platform has already demonstrated its value by winning USDOT's Intersection Safety Challenge and informing targeted local safety interventions.

CCAT, a U.S. Department of Transportation–funded University Transportation Center, helps lead the national research-to-deployment pipeline for connected and automated transportation. CCAT's core strengths include convening multi-university consortia, partnering with state and local agencies and industry, and translating research into deployable methods, tools, and workforce development, so that emerging technologies can be evaluated and implemented responsibly at scale.

Mcity, founded by UMTRI as the world's first dedicated proving ground for connected and automated vehicles, brings together controlled-environment testing, simulation at scale, and instrumented urban operations. Mcity 2.0, funded by the National Science Foundation, is a remote-accessible, mixed-reality testbed that allows researchers around the nation to conduct AV evaluations with a combination of digital twins and physical infrastructure. In 2025, Mcity also launched the **M-Air** Program, which has begun building the testbed for advanced air mobility, supporting research, testing, and safe flight operations for drones and emerging electric vertical takeoff and landing concepts.

UMTRI maintains deep partnerships with state and local agencies, industry, and standards bodies, which allows us to translate research into policy-ready guidance and deployable tools. As a public institution, our mission is to generate evidence that serves the broader public interest.