Safer Rides, Smarter Futures: How We Build Public Trust in Self-Driving Cars

By Keke Long, PhD.
Connected & Autonomous Transportation Systems (CATS) Laboratory
University of Wisconsin-Madison

Introduction and Program Acknowledgement: I am a postdoctoral researcher in the Connected & Autonomous Transportation Systems (CATS) Laboratory at the University of Wisconsin-Madison, directed by Professor Xiaopeng Li, the Harvey D. Spangler Professor. The CATS Lab is dedicated to advancing research in intelligent transportation systems, exploring how connected and autonomous vehicles can revolutionize the future of mobility and transportation systems.

I am delighted to contribute to the program "Sharing UW-Madison Postdoctoral Scholarly Research with Non-Science Audiences," sponsored by the Wisconsin Initiative for Science Literacy (WISL). This program, made possible by the dedication of the WISL staff, specifically Cayce Osborne, Elizabeth Reynolds, and Professor Bassam Shakhashiri, is instrumental in fostering connections between scientific exploration and a wider audience. My journey as a postdoctoral fellow is not just about research—it's a shared endeavor to bring the marvels of science to diverse communities. Initiatives like this are essential in making cutting-edge research accessible and relevant to everyone, helping bridge the gap between academia and the public who can benefit from and engage with scientific discoveries.

1 A Quiet Morning, An Unusual Sight

It's a crisp Saturday morning in Madison, and I'm walking to my favorite coffee shop on University Avenue. The streets are quiet—just a few joggers, some students heading to the library, and the usual weekend calm. Then I see it: a car rolling slowly down the street, its sensors spinning gently on the roof, but no one behind the wheel.

For most of us, cars have always meant having someone at the wheel. Family road trips with mom or dad driving, learning to parallel park, and the freedom of a first driver's license. Driving has been about control, about being in charge. But here, in this moment, I'm watching a vehicle navigate on its own, making decisions, stopping at crosswalks, turning corners, all without a human driver. And I can't help but wonder: Would I feel safe crossing the street in front of that car?

That question has become the heart of my research.

2 The Question Nobody Was Asking

When most people talk about self-driving cars, they focus on the passengers inside. Will the technology work? Will it be comfortable? Will people want to use it?

These are important questions. But they miss something crucial: What about everyone else?

What about the pedestrian crossing the street? The parent is watching their child play near the curb? The cyclist navigating busy intersections? The elderly neighbor who never asked for self-driving cars but now has to share the road with them?

If autonomous vehicles are going to become part of our daily lives, they need to earn trust—not just from the people inside them, but from everyone around them.

That's the gap my research aims to fill.

3 Why This Matters More Than You Think

Imagine you're a parent in a suburban neighborhood. Your kids play in the front yard, sometimes chasing a ball into the street. For years, you've taught them to watch for cars, to make eye contact with drivers, to wait for a wave before crossing.

But now, there's no driver to make eye contact with. No wave. No nod. Just a vehicle approaching—and your child standing there, uncertain.

This isn't just about technology. It's about community, safety, and how we adapt to change together.

The promise of autonomous vehicles is enormous. They could reduce the 40,000 traffic deaths that happen in the U.S. every year, 94% of which are caused by human error. They could provide mobility for people who can't drive. They could reduce emissions, ease congestion, and transform cities.

But none of that will happen if people don't trust them.

4 A Personal Journey to an Unexpected Question

I didn't set out to study trust in self-driving cars. My background is in robotics and control systems—I was fascinated by developing algorithms that could make transportation more efficient, helping autonomous vehicles navigate through traffic, optimize routes, and coordinate with each other.

For a long time, my focus was purely technical: How can we make these systems work better? How can we reduce congestion? How can we improve efficiency?

But a project during my doctoral studies changed everything.

5 The Moment That Shifted My Perspective

I was involved in an autonomous shuttle pilot program in Florida. The technology worked beautifully—the shuttle followed its routes, stopped at designated points, obeyed all traffic rules. By every technical metric, it was a success.

But what struck me most wasn't the technology. It was the people.

I remember an elderly woman who rode the shuttle regularly to visit the downtown library. She told us the shuttle gave her independence—she could now make this trip twice a week instead of once a month. The technology had genuinely improved her life.

Yet I also noticed the uncertainty in others. Pedestrians standing at crosswalks, hesitating, not knowing whether to cross. Cyclists are keeping unusually large distances behind the shuttle, unsure of its behavior.

The shuttle was doing everything correctly. But it wasn't just about being correct—it was about being *understood*. Being *predictable*. Being *trustworthy*.

That experience crystallized something for me: there's a huge gap between developing a brilliant algorithm and seeing it actually accepted in the real world. I realized I had been asking the wrong questions. Not just "Does this algorithm work?" but "Would people accept this? Would they feel safe around it? Would they trust it?"

When I joined the CATS Laboratory at UW-Madison as a postdoctoral researcher, I knew my focus had to shift. I wasn't just building algorithms anymore—I was building algorithms that communities could trust.

6 Teaching Cars to Think Like Humans

Here's the challenge: following traffic rules isn't enough.

A self-driving car can obey every law—stop at red lights, yield to pedestrians, maintain the speed limit—and still make people uncomfortable. Why? Because humans don't just follow rules. We read situations. We make judgments. We adjust our behavior based on context.

Think about it: You drive differently in a school zone at 3 PM than you do on an empty highway at midnight. You slow down when you see kids playing near the street, even if they're not in the crosswalk. You make eye contact with pedestrians to show you've seen them.

Self-driving cars need to do the same thing—not just follow rules, but understand context.

That's where my research comes in. I've developed two complementary approaches to address this challenge.

7 PERL: Teaching Cars to Handle the Unexpected

The concept came from a fundamental tension I observed in autonomous driving research. On one hand, we have solid physics models—mathematical equations that describe how cars should behave. If you press the brake pedal this much, the car should slow down at this rate.

But reality is messier than textbooks. Maybe the road is wet and the car doesn't slow down as quickly as predicted. Maybe the car ahead suddenly changes lanes. Maybe worn brake pads mean the car takes longer to stop.

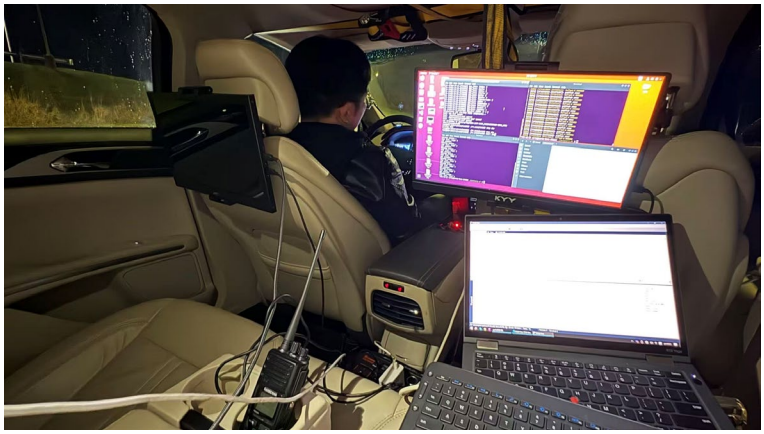
Traditional approaches force a choice: trust the physics model (and accept that it's sometimes wrong) or trust a purely data-driven AI model (and risk unpredictable behavior). I wanted both—the reliability of physics and the adaptability of AI.

The Core Idea

The breakthrough was thinking about the **residual**—the gap between what the physics model predicts and what actually happens. If the physics model says the car should slow down by 10 mph, but it only slows down by 8 mph, that 2 mph difference is the residual. Instead of ignoring these gaps or abandoning the physics model, what if we could learn to predict them?

That's exactly what PERL does. I combined a physics-based model with a neural network that learns to correct the physics model's predictions in real-time. The physics keeps the system grounded—the car can't violate the laws of motion. But the neural network fills in the gaps, handling all the messy, unpredictable stuff that makes real driving so complex.

Testing on Real Vehicles



Collect data in our CATS Lab autonomous vehicle.

After developing PERL in simulation, I tested it on our lab's autonomous vehicle. The results confirmed what I hoped: by combining physics-based reliability with AI's adaptability, the system could handle real-world variations smoothly. When encountering a wet road or unexpected traffic patterns, PERL didn't just react—it adapted, maintaining the kind of smooth, predictable behavior that makes people feel comfortable.

But PERL also taught me something that would shape my next project. While the algorithm successfully combined physics and learning to handle vehicle dynamics, I realized there was another gap: understanding *context*. PERL could adapt to how the car physically responds, but what about understanding *why* a pedestrian is standing at the curb? Or *what* those construction cones ahead actually mean?

That realization led me to my next algorithm.

8 VLM-MPC: Teaching Cars to Understand, Not Just See

PERL solved one problem—helping cars adapt to physical variations in the real world. But it couldn't solve everything.

Imagine you're driving through a neighborhood and see a ball roll into the street. Your brain immediately thinks: "A ball means a child might be chasing it. I should slow down, even if I don't see the child yet."

That's not just seeing—that's understanding. It's using reasoning to predict what might happen next.

Traditional autonomous driving systems, including PERL, are good at reacting to what they detect: "There's a car. There's a pedestrian. There's a stop sign." But they struggle with context and reasoning: "Why is that pedestrian standing at the curb? What are they likely to do next? What do those construction cones actually mean for how I should drive?"

This limitation bothered me, especially as I watched the field evolve. Around this time, large language models like ChatGPT were making headlines for their ability to understand and reason about complex situations. Some researchers started proposing: why not use these powerful AI models to control autonomous vehicles directly?

The Problem I Saw

But I was skeptical. Having worked extensively with AI models through PERL, I understood their limitations. Large language models, for all their impressive capabilities, have a fundamental problem: they hallucinate. Even when generating simple text, they sometimes make errors or produce unreliable outputs.

If GPT can occasionally give you a wrong answer when you ask it a question, how could we trust it to control a vehicle traveling at 30 mph down a city street? The stakes are too high. A single moment of "hallucination" in vehicle control could be catastrophic.

Yet these models clearly had something valuable—their ability to understand context, to reason about situations, to interpret complex scenes the way humans do.

The Solution: Combining Strengths

This is where my experience with PERL became crucial. PERL had taught me the power of combining different approaches—letting each component do what it does best. The physics model provided reliability; the neural network provided adaptability.

Could I apply the same principle here?

That insight led me to develop VLM-MPC—Vision-Language-Model-based Model Predictive Control.

The key innovation was recognizing that we don't need the language model to control the vehicle. We just need it to understand the scene. Let the AI do what it's good at—reasoning and interpretation. Then hand that understanding to a proven control system that can make reliable, safe decisions.

Here's how it works: The vision-language model acts as the "eyes and brain" of the system. It looks at the road ahead and interprets what it sees:

- It sees construction cones and understands: "This means the normal traffic pattern is disrupted. Proceed with extra caution. Workers might be present."
- It sees children playing on a lawn near the street and reasons: "Children are unpredictable. They might run into the road. The vehicle should slow down preemptively."
- It notices a pedestrian standing at a crosswalk looking at their phone and infers: "This person might not be paying attention. Be prepared to stop even if they don't seem ready to cross."

But here's the critical part: the language model doesn't control the car. It simply provides a high-level understanding—what the situation is, what might happen, what caution level is needed. This understanding is then passed to a traditional Model Predictive Control (MPC) system, which translates those insights into precise, safe vehicle commands.

The MPC handles the actual control—how much to brake, when to steer, maintaining safe distances. This is proven technology, mathematically guaranteed to be safe within its operating constraints. No hallucination, no uncertainty in the control decisions.

Developing and Testing VLM-MPC

I developed and tested VLM-MPC entirely within virtual environment. VLM-MPC represents a more experimental approach, combining cutting-edge AI with traditional control. Testing it thoroughly in simulation first allows me to explore its potential while ensuring safety.

In the simulation, I tested VLM-MPC across diverse driving scenarios. The results were promising. The system demonstrated something closer to human-like reasoning. It didn't just react to objects in its path—it anticipated potential risks based on context. It slowed down near playgrounds even when no children were visible. It gave extra space to cyclists even before they signaled a lane change. It recognized that a stopped school bus meant children might be crossing, even if they weren't yet in the crosswalk.

9 Two Algorithms, One Philosophy

Looking back at PERL and VLM-MPC together, I see a consistent philosophy in my research: don't force a single approach to do everything. Instead, combine different methods, letting each handle what it does best.

PERL combines physics and neural networks. VLM-MPC combines language model reasoning and traditional control. Both recognize that the future of trustworthy autonomous vehicles isn't about finding one perfect technology—it's about intelligently integrating multiple approaches.

And both were designed with the same ultimate goal: creating autonomous vehicles that people will trust, not just because they're safe, but because they behave in ways that make sense to humans.

10 Building a Digital Twin of Madison

To develop and test trustworthy autonomous driving algorithms, I needed a safe space to experiment—somewhere I could test thousands of scenarios without any real-world risk. That's where my virtual city comes in.

I started by collecting real data. I flew drones over three key intersections in Madison, capturing hours of traffic footage. From these videos, I extracted real driving parameters: how aggressively do Madison drivers typically accelerate? How much space do they leave when following another car?

I then used simulation software to build a digital replica of these intersections, programming the virtual vehicles to behave like real Madison drivers using the parameters I'd extracted.

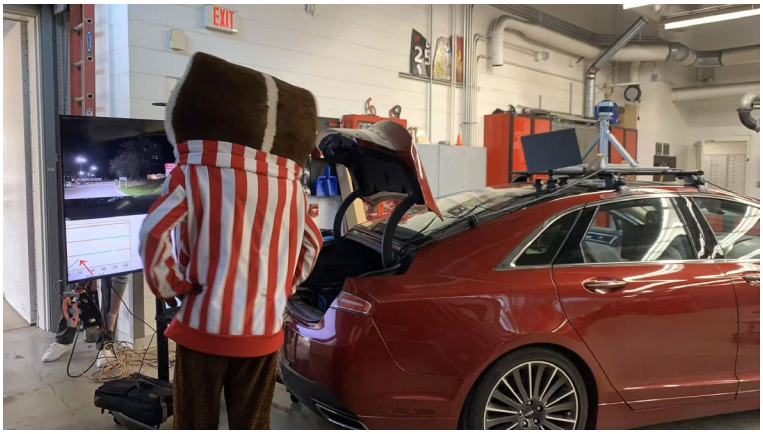
The beauty of this virtual environment is its flexibility. I can simulate a sunny afternoon with light traffic, then immediately switch to a rainy rush hour with aggressive drivers. I can place a pedestrian at a crosswalk, a cyclist in the bike lane, a child playing near the curb—and see how my autonomous vehicle algorithms respond.

This virtual testing environment has been essential for both PERL and VLM-MPC. It's where I validate that these algorithms don't just work in theory, but can handle the messy, unpredictable reality of urban driving.

11 The Bigger Picture

Self-driving cars are coming. But their success won't be determined by engineers alone. Parents will decide whether it's safe for their kids to play outside. Cyclists will decide whether they feel comfortable sharing bike lanes. City planners will decide whether to allow these vehicles in residential areas.

Technology adoption isn't just a technical problem. It's a social one, and autonomous vehicles need to earn trust through integration into our communities.



Welcome Bucky to check our autonomous vehicle and devices

12 What Happens Next

My research continues to evolve. Currently, I'm focusing on improving the stability and reliability of large language models in autonomous driving contexts. The goal is to make systems like VLM-MPC even more robust and dependable.

Beyond autonomous vehicles, I hope the approach I've developed—combining AI reasoning with proven control systems—can help large language models find broader and safer applications in other domains where reliability is critical.

I'm also exploring new questions: How can self-driving cars communicate their intentions to pedestrians? How do different cultures perceive autonomous vehicles differently? How do we maintain trust after an accident?

13 A Vision for the Future

Imagine a future where self-driving cars are so well-integrated, so trusted, that we don't think twice about them. Where children play safely near streets because cars are programmed to be extra cautious around them. Where elderly people regain independence through autonomous shuttles. Where traffic accidents become rare because human error is largely eliminated.

That future is possible. But it won't happen through technology alone.

It will happen through research that prioritizes human needs. Through design that considers everyone affected. Through testing that goes beyond technical metrics to measure comfort, trust, and acceptance.

It will happen when we stop asking "Can this work?" and start asking "Will people welcome this?"

14 Back to That Quiet Street

So, back to that Saturday morning in Madison. A driverless car rolls by. A child is playing nearby.

Would I feel safe?

After years of research, I can say: it depends. It depends on whether that car was designed with trust in mind. Whether it was tested for smoothness and predictability, not just safety. Whether it behaves in ways that make sense to humans, not just algorithms.

The good news? I'm working on it. Every simulation, every algorithm refinement, every conversation with community members brings us closer.

The future of transportation isn't just about autonomous vehicles. It's about autonomous vehicles that we trust, that we welcome, that we feel safe around.

And that's a future worth building.