Pravni fakultet u Rijeci

Who Do We Trust When No One Is Driving? Rethinking Trustworthiness in Automated Vehicles

Ivana Kunda

The 9th International Scientific Conference Petar Šarčević "Automated Vehicles at the Crossroads: Law, Policy and Deployment in Transition"
13 and 14 November 2025, Faculty of Law in Rijeka

The Trustworthy and Explainable AI Lab
The Center for AI and Cybersecurity
Sveučilište u Rijeci
University of Rijeka

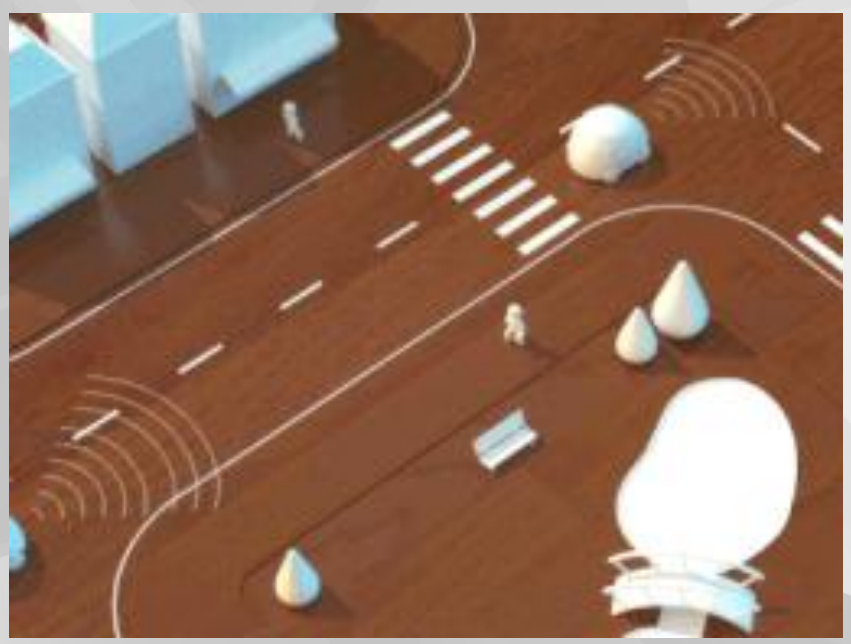
Funded by the European Union
NextGenerationEU

Trustworthiness

- **being** trusted – **deserving** trust – **earning** trust

Domain	Definition	Sources
Governance	"Lawful, Ethical, Robust"; fairness, accountability, transparency	EU HLEG (2019); Jobin et al. (2019)
Technical	Robustness, calibration, differential privacy, fairness metrics	Szegedy et al. (2014); Guo et al. (2017)
HCI	Predictability, competence, understandability, trust calibration	Lee & See (2004)
Philosophy / STS	Moral trustworthiness, institutional legitimacy, autonomy	O'Neill (2018); Coeckelbergh (2020)

Trustworthiness in the AV – 101



- safety validations
- sensor reliability
- failover modes
- explainability of models
- regulatory compliance

Relations, not just statistics

Trustworthiness depends on:

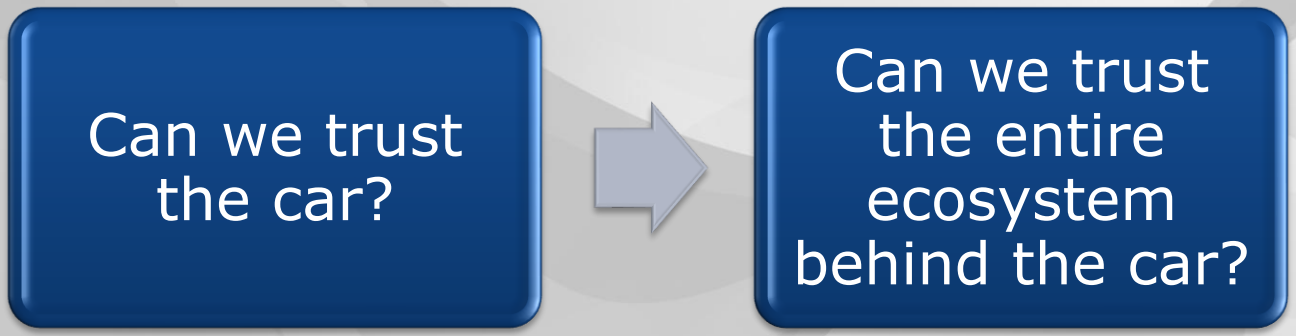
- how humans **interact** with the AV
- how predictable and understandable the AV **feels** to people
- how the AV communicates **intentions**
- how **responsibility** is allocated

People may mistrust AVs even when statistically safer than human drivers because **trust is experiential**, not merely probabilistic.

Distributed responsibility, not just technical improvments

Trustworthiness increases with:

- explicit responsibilities for failures and updates
- transparent lines of accountability
- users know who is responsible for element designs, maintenance, updates, decision logs



Delivering the promise, not just piles of paperwork

- how AV behaves in everyday traffic, not just technical documentation
- user experience (smooth braking, understandable handovers) matters more than abstract certifications
- public road trials, transparency reports, and near-miss disclosure

Communication, not just fixing code

AVs should communicate intentions clearly to:

- passengers
- pedestrians
- other drivers

Examples:

- external displays that signal yielding
- predictable driving patterns that “feel” human-friendly
- interfaces that explain what the car is perceiving

Social legitimacy, not just technical assurance

Because AVs reshape public space, we need:

- public participation
- democratic governance
- inclusive decision-making

Trustworthiness involves:

- community engagement in deployment
- transparency in safety demonstrations
- consideration of vulnerable groups (impaired, elderly, children, consumers, migrants)

Conclusions

Not only

- accurate perception
- reliable control algorithms
- robust safety validations

Trustworthy AV system is embedded in trustworthy human, organizational and societal **relationships**

But also

- meaningful and predictable **interactions** with humans
- trustworthy **institutions** behind development, testing, and deployment
- socially embedded **communication** and transparency practices
- distributed **accountability** and responsible governance
- cultural acceptance and public **legitimacy**

Your questions or comments are also welcome at:

ivana.kunda@uniri.hr

Thank you for the attention!

Legal disclamer:

This paper has been written and presented within the scope of the project „TECHNO-IUS: Legal and Ethical Implications of the Dehumanisation of the Justice and Public Administration through Legal Technology" (uniri-mzi-25-31), funded by European Union – NextGenerationEU via the Croatian National Recovery and Resilience Plan 2021-2026, in conjunction with the University of Rijeka, Faculty of Law Programme Financing. The views and opinions expressed are solely those of the authors and do not necessarily reflect the official position of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.