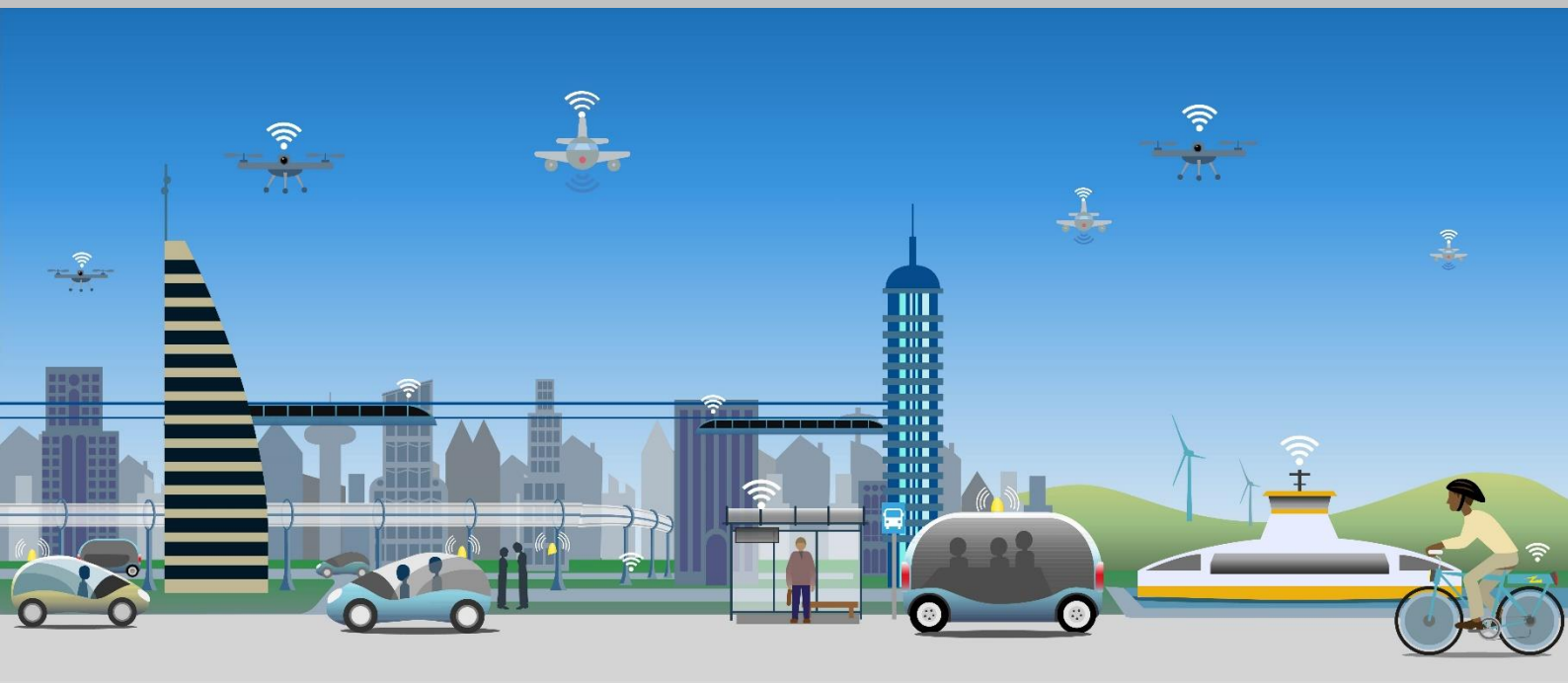
Verification of Autonomous Mobile Systems: Directions for Future Developments

Michael Fisher, Javier Ibanez-Guzman, Abdelkrim Doufene, Karla Quintero

POLICY REPORT

Verification of Autonomous Mobile Systems:
Directions for Future Developments



Recommendations from the *First International Symposium on the Verification of Autonomous Mobile Systems, March 2023*

Report created by:
Michael Fisher (University of Manchester), Javier Ibanez-Guzman (Renault Group),
Abdelkrim Doufene (IRT SystemX), Karla Quintero (IRT SystemX)
September 2023

Symposium supported by:

IEEE Robotics & Automation Society
Institut Mines-Télécom
MANCHESTER 1824 The University of Manchester
NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE CETRAN
Renault Group
SystemX INSTITUT DE RECHERCHE TECHNOLOGIQUE

BACKGROUND

Mobility is a basic daily necessity that is undergoing a major transformation driven both by societal needs and by technological developments. There are requirements to provide broader accessibility, to contribute to sustainable mobility solutions, to improve safety, etc. In addition, progress in a range of areas, such as sensing, algorithms for perception, verifiable decision-making, computing power, and communications technologies, together with more efficient machine learning methods, are enabling the development of autonomous systems capable of interacting with complex and dynamic environments.

Over recent years, we have witnessed an increasing interest in the deployment of autonomous vehicles across land, sea and air demonstrating their potential to transport people and goods as well as providing solutions to a wide range of field applications. However, mobile autonomous systems with high levels of autonomy remain scarce, with only a few and each predominantly operating under very constrained conditions or relying on specialist operators for safety assurance. Deploying truly autonomous systems has been found to be more difficult than expected, for example, autonomous passenger vehicles.

The acceptability of autonomous mobile platforms depends on our ability to demonstrate their safe, reliable, and continuous operation. One of the major concerns is safety when operating in open or public spaces; providing evidence to all stake-holders of the safe behaviour of mobile systems in these situations remains a challenge. This is compounded by the need to interact with other traffic agents whose behaviour might well be unpredictable. Experience has shown that there can be a very large number of scenarios that these systems need to address even in well-defined operational design domains. Extensive testing in real conditions, though enabling some progress, will be insufficient to address all these situations. Meanwhile, society might be particularly unforgiving of failures, compromising long term progress across the area.

Considerable work has been carried out, both in industry and academia, towards demonstrating how assurance in such domains might be achieved, as we strive to deploy autonomous mobile systems in public spaces independent of human control. This includes the development of related standards. However, there clearly remain significant steps that must be taken. Some of these may concern research that is needed, some concern practical development that needs to be undertaken, some concern publicising existing solutions and educating a range of stakeholders, etc. But what ***are*** these steps? Where are the ***gaps***? And which developments are the most ***important***?

Disclaimer

This report summarizes feedback from experts in various sectors as they provide assessments based on available evidence and relevant standards in their fields of expertise. Their contributions do not represent the general strategy or positioning of their institutions. Furthermore, they do not constitute legal advice and the contributors of the report disclaim any responsibility for how this report is used or interpreted beyond its intended scope.

SYMPOSIUM

This symposium gathered together experts from across relevant domains with participants including practitioners, regulators, and scientists working in domains relevant to, or specific to, autonomous mobile systems. The purpose of the symposium was to combine different points of view in order to enrich the knowledge of this emergent community, creating awareness of the available scientific methods whilst benefitting from experience acquired by industry in their latest developments and viewpoints, and from regulators tasked with overseeing these new technologies. Multiple questions remain around our ability to embed verifiable properties as part of the design of autonomous systems, on the methods for their verification, on the combination of mathematical methods and field testing, and on addressing the complexities brought by the introduction of data intensive methods (e.g. machine learning) that introduce a further dimension of uncertainty.

The symposium was supported by the Institute for Technological Research (IRT) SystemX[1], the University of Manchester[2], Renault Group[3], Institut Mines-Télécom, the IEEE Technical Committee on the Verification of Autonomous Systems of the Robotics & Automation Society[4], and the Nanyang Technological University Centre of Excellence for Testing & Research of Autonomous Vehicles [CETRAN][5]. The event was co-chaired by Abdelkrim Doufene [Director of Strategy and Programs, IRT SystemX, France], Michael Fisher [Royal Academy of Engineering Chair in Emerging Technologies, University of Manchester, United Kingdom]; and Javier Ibanez-Guzman [Autonomous Systems Corporate Expert, Advanced Engineering, Renault Group, France] and was organised by the committee described at the end of this document[6].

FINDING ANSWERS

While dissemination and collaboration were important elements of the symposium, the core aspect involved participants working together to answer questions related to "what are the gaps?" and "which are the most important?". Consequently, the symposium involved two breakout sessions.

Session 1: What are the key issues? Why are autonomous vehicles/systems not widespread?

[The purpose of this session was to gain a common understanding of the barriers that have emerged from current efforts.]

Are the barriers **Technical**?

- Are some key R&D breakthroughs still needed? If so, where?
- Perception? Architectures? V&V? Decision-making? XAI? …

Are the barriers **Commercial**?

- Is there no business case for these systems?

Are the barriers **Regulatory**?

- Are the regulations not in place to allow them? Lack of Standards?

Are the barriers related to **Public/Governments**?

- Do the public not yet have an appetite for the technology?
- Do the public/governments not yet trust the technology?

Are these issues the same for each air/land/sea sector and also the same whether viewed from Industry/Academic/Regulatory/Public bodies? Does the local operating domain add complexity?

Session 2: Where are the solutions? What are the priorities?

[The purpose of this session was to brainstorm searching for different solutions and priorities benefitting from the experienced gathering.]

Where should we ***target*** our efforts?

- Technical issues? If so, which ones?
- Commercial issues? If so, how?
- Regulation and Standards? If so, where?
- Public awareness/appetite/trust?
- Government? How can we change views?
- Others?

What are the priorities? What are the key bottlenecks?
Are the issues/priorities the same for autonomous mobility across sectors?

---

[1] https://www.irt-systemx.fr/en/

[2] http://www.manchester.ac.uk

[3] https://www.renaultgroup.com/

[4] https://www.ieee-ras.org/verification-of-autonomous-systems/

[5] https://cetran.sg/

[6] Held in Saclay on 9th/10th March 2023: https://www.irt-systemx.fr/en/evenements/vams-is-23 Although the symposium was initially intended to be purely a face-to-face event, social unrest across France coincided with the symposium and meant that several participants could only attend virtually.

Are the priorities the same for Industry/Academic/Regulators/Public/etc.?

PRESENTATIONS

Excellent presentations, across a range of issues, were provided by the speakers.

There were talks concerning cross-cutting issues such as modelling/validation [Joseph Sifakis], architectures [Charles Lesire], verification of decision-making [Louise Dennis], model-driven development [Ansgar Radermacher], simulation-based testing [Justin Dauwels], certification of machine learning [Xiaowei Huang], computational complexity [Bruno Monsuez] and risk-based approaches [Subramanian Ramamoorthy]. Cross-cutting issues, such as safety/security [Antoine Rauzy], were explored as were a range of sector-specific aspects: autonomous rail vehicles [Christian Schindler] and their regulation [Vaibhav Puri], the regulation of autonomous ocean vessels [Ruth Taylor], verification of air mobility [Jean Daniel Sülberg], road vehicle analysis [Andrea Leitner], and standards for road vehicle autonomy [Simon Rößner].

Several posters were also presented during the breaks.

TC-VAS RECAP OF VERIFICATION TECHNIQUES[7]

There are many verification techniques aimed at assessing some ***Requirements*** of the ***System*** in question, summarised as follows.

F: Formal Verification Techniques

F1: Proof - where the behaviour of the System is described by a logical formula and a proof is carried out to establish that this logical formula implies the Requirement.

F2: Model Checking - where the Requirement is exhaustively assessed against a representation/model of all possible execution paths of the System. Program Model Checking is a variation where the Requirement is assessed against all possible execution paths of the System.

F3: Runtime Verification (sometimes referred to as *dynamic fault monitoring*) - where the Requirement is assessed against the System as it is executing.

F4: Probabilistic Verification - all the formal verification options above have probabilistic versions.

E: Empirical Verification Techniques

E1: Software Testing - where the Requirement is checked on a subset of the possible executions of the System.

E2: Simulation-based Testing - similar to above where a simulation of the real environment is used for environmental interactions.

E3: Observational – assessing the long-term, practical operation of the System.

E4: Practical Experiments – where the Requirement is assessed of the real/full system in place.

[7] As developed by the IEEE Technical Committee on the Verification of Autonomous Systems.

WHAT ARE THE ISSUES?

Technical

- *Simulations.* High-fidelity models are needed but there is still a significant (and sometimes unknown) gap between simulation and reality.
- *Scenarios*. We still do not have a clear and comprehensive set of scenarios specified for our autonomous mobile systems. We need to capture the complexity and multi-dimensionality. To what extent are the autonomous vehicle issues operational domain-specific?
- *Specification*. We (developers/regulators/public) must be clear about the capabilities and limitations of both the system and its components. We must also be clear about the interactions between systems and between humans and the system.
- *Architectures*. Many systems are built in an ad-hoc/opaque way. The key components should be identifiable and transparent (and hence verifiable and explainable). There is currently a lack of principled design approach for the whole system.
- *Verification*. There is a tension between design methods and operational evaluation - different verification and validation methods are used in each. General lack of strong formal/mathematical models across the variety of components.
- *Data-Insensitivity*. Along with the potential over-use of data-intensive components such as Machine Learning, there is often false confidence in "the more data we collect, the more accurate we are". In contrast, there are other situations where insufficient real data exists.
- *Scalability*. Although strong verification methods, such as formal verification, might work for small/select components, there remains a significant issue of scalability.
- *Runtime*. Sometimes false confidence in runtime verification emerges. For example, constructing the system in an ad-hoc way and then adding a range of runtime verification components checking for safety/regulation violations. This can be useful but is not a general solution to all the issues.
- *Sensors*. Sensing and recognition remain difficult. How much of the (road) infrastructure can/should be used to help with situational awareness? Who will pay for this?
- *Safety*. What is the degree of readiness for deployment of these technologies in public spaces?
- *Diversity*. Need to address the human diversity and local issues across a range of deployments and scenarios. Autonomous mobile systems must not just be safe for the tested scenarios.
- *Uncertainty.* Numerous sources of measurement uncertainty exist, and this propagates through the whole system. How can we measure this, and can we bound the uncertainty?
- *Autonomy*. What is special about autonomy? If there is nothing new, then we just use traditional techniques. But if there is something new (e.g. autonomous decision-making) then we need to develop and apply new design/verification/assurance techniques.

Societal/Public

- *Overselling.* Too many expectations from the public (and regulators), often due to over-hyping by researchers/manufacturers.
- *Motivation*. What is the motivation for wanting/needing autonomous mobility? And will the consumer expect longevity from vehicles?
- *Acceptance*. Different society stakeholders must accept autonomous systems for them to be deployable.

Industry/Business

- *Business Model.* What is the business model for autonomous vehicles in particular? Specialist? General consumer products? Are they too expensive?? What is the value for the customer? For example, the emphasis of autonomous vehicle development appears to shifting from the luxury market to public/mass transportation.
- *Mobility as a Service*. If autonomous vehicles exist, why should consumers buy other types of vehicles?

Regulatory

- *Methods.* No clear method to validate autonomous mobility/vehicles. Consequently, no risk acceptance and so regulations are often conservative. Vehicle manufacturers might take an economic/liability approach.
- *Sector Norms*. Many areas (e.g trains) are highly regulated, with homogenous views. Steps towards autonomy are slow and long: safety driver/attendant; etc. Several norms on autonomous driving systems with high levels of autonomy and the introduction of AI-based components under development are demonstrating the need for precise definitions and verification methods.
- *Evidence*. What kind of evidence do regulators need? And is this the same as the evidence developers need/provide? And will insurers/public need different/stronger evidence?
- *Knowledge*. Emphasize the need for domain/technological knowledge within the regulator.
- *Evolution*. No obvious timeline to implement norms in regulators across sectors.
- *Jurisdictions*. Regulations are being developed in very different ways across the world. Regulations being developed in the USA tend towards certification and are more flexible. Regulations being developed in the EU/UK are more rigid and require developers to make the safety case.
- *Liability*. Who will handle this? Insurance companies? Manufacturers? Governments? And will this be acceptable to the public?

WHAT ARE THE PRIORITIES?

Note that, while all the items mentioned below are important and require exploration, we have made an effort to prioritize these into three categories.

High Priority

- *System Architectures.* Rather than concentrating on individual technologies (perception, navigation, decision-making, communication, etc.), it is important to consider how all the elements work together and to understand the differing criticalities and diverse verification techniques that will be significant to different parts of the system. Architectural aspects, such as modularity and compositionality, support transparency (and hence strong verification), encapsulation of functionality, explainability, etc, and this *holistic development* that will allow us to devise more transparent, more flexible, more resilient, and more verifiable systems.
- *Precise Specification.* We must be clear about exactly what the system, and its sub-components, should be doing. Vague descriptions such as “it will be safe” are not helpful. Part of this clarity in specification also covers the description of edge-case behaviours - what should the system or component do in unusual situations? This is important, not only for verification (since we must know what to verify) but also for resilience, transparency, etc.
- *Take Small Steps to Autonomy*. In particular, verify in depth the system before moving on to the next (small) step. Although the move from human-controlled to autonomous vehicles is not continuous, we should move through step-changes in small increments, where possible. We should advance directly from structured/controlled environments to public spaces. This approach is not only a benefit for verification but also helps in understanding and assessing system resilience.

Medium Priority

- *Education of Engineers.* Especially with new technologies, such as autonomous systems, machine learning, etc., there is a tendency to either assume everything is new and must be re-invented or assume there are no engineering processes to follow and produce systems in an ad-hoc manner (and sometimes both). It is important that developers (and possibly regulators) be educated about standard software engineering processes, such as clear specification (as above), modularity, traceability, the benefits, and drawbacks, of different verification approaches, system functional architectures, etc. Anecdotally, many of the failures in autonomous mobile systems are due to poor systems development practices and “cutting corners”.
- *Heterogeneous Approaches*. Use the best form of verification in the appropriate situation. Using multiple different verification techniques both during design (formal, simulation) and operation (runtime verification), and combining them together, can help confidence in the overall system.
- *Clarity of Language.* As with the clarity of specification above, it is important that we are clear with what we mean. This especially occurs with autonomy and the role of human operators/drivers. We must understand, and if necessary, explain, what we mean by “autonomous”, “automated”, “adaptive”, “semi-autonomous”, “remote pilot”, “intelligent”, etc., rather than using these terms in an ad-hoc manner. Since we all may have our own interpretations of these terms it is important that we are clear and consistent. Similarly, with language concerning mobile activity: “near miss”; “didn’t see”; “avoid”; etc.
- *Decision-making*. Understanding, explaining, and verifying decision-making is crucial to autonomous system deployment. An autonomous system will make its own decisions and so we need to be very clear about when, and how, these decisions are made especially in scenarios we have not predicted.
- *Public/Government Education/Expectation Management*. Over-selling of the technology can set unrealistic expectations. Then, when not met we tend to “move the goalposts” and change what we promised. This is related to the “clarity of language” issues above in that we might promise a “driverless” car but eventually, we market a car where the responsibility still mainly lies with the driver. Similarly, we must be clear that no technology is guaranteed to be 100% safe and so some *risk* will be involved. It is necessary to educate all stakeholders.

Lower Priority

- *Target "High Benefit "Problems.* Related to issues with business models for autonomous mobility, it is important that researchers target those systems that will be of demonstrably high benefit. Just providing an autonomously mobile solution for tasks that could easily be achieved with less advanced technologies does not help the field.
- *Better Economic/Business/Legal Ecosystem.* Not a technological issue, but we need to ensure that the ecosystem support for autonomous mobility is in place. This includes the supply chain, the business models, the regulatory frameworks, the insurance mechanisms, etc. And, relevant to verification, these elements must be sensitive to the different levels of verification provided.
- *Better Software Engineering.* While we above cited the need to utilise traditional software engineering where possible, it is also important that these software engineering methods be developed. The traditional methods do not consider the very different aspects of autonomous systems and AI components and, until they do, several dimensions of software engineering will remain missing.

CONCLUSIONS

This symposium aimed to promote the transfer of knowledge and experience between those involved in the design, development and governance of autonomous mobile systems operating over large outdoor spaces and to generate priorities relevant to the diverse organisations involved in the use of mobile autonomous systems across land, air, and sea applications.

Autonomous mobility issues are (at a high-level) general across air, land, and sea, and many common issues emerge. This report highlights several gaps that prevent us from deploying reliable, trustworthy, and safe autonomous mobility in unconfined spaces. The issues highlighted span technical, regulatory, business, and governmental contexts.

The introduction of AI-based methods (such as machine learning) has enabled important progress to be made across autonomous mobile systems, whilst at the same time adding significant additional complexity. These methods can introduce coding errors that are difficult to trace, failure in these components is often difficult to predict, generalisation is a significant challenge, there is a lack of transparency, and consequently explainability and verifiability, of decisions made, etc. Consequently, degrees of uncertainty introduced with these approaches adds much complexity to the verification of autonomous systems.

We believe that verification of autonomous systems should provide the theory, methods, metrics, and processes that facilitate the deployment of autonomous mobility. We need a holistic approach where major stakeholders can participate, so autonomous systems can be accepted by society. By tackling the issues highlighted in this report, we believe that significant progress can be made across autonomous mobile systems in general.

PARTICIPANTS [Org = Organising Committee]

Abdelkrim Doufene	IRT SystemX, FR	[Co-Chair; Org]
Ana Isabel Garcia Guerra	Nanyang Technological University, NTU, SG	[Org]
Andrea Leitner	AVL, AT	[Speaker]
Ansgar Radermacher	CEA, FR	[Speaker]
Antoine Farha	AVL, FR	
Antoine Rauzy	Norwegian University of Science and Technology, NO	[Speaker]
Arunkumar Ramaswamy	Renault Group, FR	
Ben Hiett	Defence Science and Technology Laboratory, UK	
Bruno Monsuez	ENSTA Paris-Tech, FR	[Speaker]
Charles Lesire	ONERA, FR	[Speaker]
Christian Schindler	Aachen University, DE	[Speaker]
Christophe Bohn	IRT SystemX, FR	
Clément Zinoune	Renault Group, FR	
Daniele Sportillo	MathWorks, FR	
Darryl Hond	Thales Research, Technology and Innovation, UK	[Virtual]
Dejanira Araiza-Illan	Johnson & Johnson Ltd, SG	[Virtual; Org]
Emmanuel Alao	Université de Technologie de Compiègne, UTC, FR	
Federico Camarda	Renault Group, FR	
Gabriele Incorvaia	Thales, UK	
Hamid Asgari	Thales Research Technology and Innovation, UK	
Hatem Hajri	Safran Group, FR	
Imane Taourarti	Renault Group, FR	
Javier Ibanez-Guzman	Renault Group, FR	[Co-Chair; Org]
Jean Daniel Sülberg	German Aerospace Center, DE	[Speaker]
Jing Yew Yap	Nanyang Technological University, NTU / CETRAN, SG	
Joseph Sifakis	Université Grenoble-Alpes, FR	[Virtual; Speaker]
Justin Dauwels	TU Delft, NL	[Speaker]
Karla Quintero	IRT SystemX, FR	[Org]
Kevin Leahy	MIT Lincoln Laboratory, US	[Virtual]
Louise Dennis	University of Manchester, UK	[Speaker]
Lounis Adouane	Université de Technologie de Compiègne, UTC, FR	
lyes Saidi	Université de Technologie de Compiègne, UTC, FR	
Michael Fisher	University of Manchester, UK	[Co-Chair; Org]
Michel Batteux	IRT SystemX, FR	[Org]
Mohamed Tlig	IRT SystemX, FR	[Org]
Morayo Adedjouma	CEA List, FR	
Niels de Boer	Nanyang Technological University, NTU, SG	
Paul Labrogère	IRT SystemX, FR	
Peter Boyd	Defence Science and Technology Laboratory, UK	
Philippe Xu	Heudiasyc, UTC/CNRS, FR	
Rhandy Cardenas	Université de Technologie de Compiègne, UTC, FR	
Ruth Taylor	UK Maritime and Coastguard Agency, UK	[Speaker]
Signe Redfield	Naval Research Lab, US	[Virtual; Org]
Simon Rößner	TÜV Süd Auto-Service GmbH, DE	[Speaker]
Subramanian Ramamoorthy	University of Edinburgh, UK	[Speaker]
Thibault Charmet	Renault Group, FR	
Thomas Chauviere	SYSTRA, FR	
Vaibhav Puri	Rail Safety and Standards Board, UK	[Virtual; Speaker]
Vincent Honnet	IRT SystemX, FR	
Wen-Hua Chen	Loughborough University, UK	
Xiaowei Huang	University of Liverpool, UK	[Virtual; Speaker]