

Read in the Engineer [Here](#)

Comment: Automotive and defence can learn from each other in the quest to define an ethical framework for automated vehicles

Dr. Helen Monkhouse, Chief Engineer for Functional Safety at HORIBA MIRA, explores the comparison between automated vehicle development in automotive and defence applications



One of the biggest ethical questions facing automated vehicle developers is that of responsibility and accountability. With the growing use of artificial intelligence (AI), the design team potentially has less control over the operational actions taken by an automated system, and less visibility of the decision making behind those actions.

The void between action and action responsibility is what philosophers refer to as a 'moral responsibility gap'. The widening of this moral responsibility gap is a challenge shared by both civil and defence industries. It makes it increasingly difficult to define who is ultimately responsible for a vehicle's actions, and to apportion blame if things go wrong.

This question of responsibility and blame potentially means different things to different sectors. Military vehicles, for instance, face a very different operational design domain to their civilian counterparts. The defence industry also has different approaches to safety assurance, not to mention very different timescales and commercial pressures. So, could the engineers working in these two sectors learn from each other?

To some extent, all vehicles will encounter scenarios that haven't been explicitly tested in their development. Vehicle collisions in the real world, for instance, very rarely occur at the precise speed and angle prescribed by laboratory crash tests. **However, these tests are sufficiently representative that manufacturers can be confident that the results will extrapolate to the vehicle's real-world performance, ensuring that it meets or exceeds the standards laid down.**

How do we protect the engineers who now find themselves with the responsibility for developing and releasing something that can never be fully analysed or fully tested

In contrast, the sheer complexity of automated driving – an almost infinite range of different road layouts, weather conditions and traffic behaviours with potentially no human oversight – means that there is a far greater chance of encountering a scenario that has not been validated, even with the most diligent and comprehensive test programme.

So, where does the responsibility for the actions that an automated vehicle takes in these unknown scenarios lie? And how do we protect the engineers who now find themselves with the responsibility for developing and releasing something that can never be fully analysed or fully tested?

The same questions apply in military applications, but the context is very different. Leaving aside the application of lethal force, where human oversight will be retained for the foreseeable future, what similarities and differences exist for automated vehicles operating in civilian and defence contexts?

Driverless robotaxis are already operating in the civilian world. As with the partially-automated driving functions now offered on a lot of mainstream vehicles, these are expected to be operated without any significant training. In contrast, the personnel using and interacting with military vehicles will typically have received extensive training. Furthermore, while military vehicles may face hostile action, the fundamental environment that they navigate and operate within will typically be simpler than a bustling city full of vehicles, cyclists and pedestrians.

To make this chaotic environment more manageable, civilian transport is governed by clear traffic rules, and the emphasis is on taking sufficient time to make safe and accurate decisions. A military vehicle, on the other hand, may need to make split-second decisions to respond to a potential threat or gather time-sensitive intelligence. Here, the focus is on speed and adaptability.

Military organisations have their own legal and ethical obligations, but they tend to be less visible and less sensitive to public opinion. In a civilian context, where private car ownership is high, building and retaining consumer confidence and brand image is paramount to success. Therefore, it is perhaps no coincidence that commercially risky self-driving development projects are typically divorced from the parent automotive OEM brands.

As a society, we accept that humans are fallible and even the best drivers on occasions will make mistakes and have accidents, but we struggle to apply the same logic to machines. That begs the question of where we set the threshold for safe operation and, as a society, how many machine induced accidents are we willing to accept?

Compounding this question is the fact that humans, as a society, tend to be ill-equipped to reason rationally about the risks and rewards associated with individual endeavours. Hence, with such high levels of uncertainty at play, when can the developers be judged to have discharged their responsibilities? Perhaps the most logical argument is that an automated vehicle only needs to be safer than the average careful and competent human driver, but will society accept that position?

Again, context is key. If the use of AI in a military application can be shown to save lives on the battlefield, it's likely to be accepted, even if there are still risks attached. In a civilian context, for the individual suffering a heart attack, who chooses to take a robotaxi to the hospital rather than waiting for an ambulance, the risks associated with the automated vehicle operation might be justified. However, just one high-profile failure could be enough to turn public opinion against the project.

There is also a cultural divide between the two industries. Automotive tends to push through evolutionary changes more frequently, but in smaller steps. Conversely, the underlying vehicle platform can remain virtually unchanged for decades in defence, albeit with a variety of engines and weapons systems applied during that time. Perhaps the most famous example of this is the Boeing B-52, which first flew more than 70 years ago, and is widely tipped to be the first aircraft to see 100 years in continuous military service.

In automotive, safety testing historically has been performance-led; if a system demonstrates the required capabilities during conformity testing, the vehicle is approved for sale. For highly regulated industries, such as aerospace, the approval process places far more scrutiny on the engineering **procedures** used to achieve those results, as well as how they will be updated and maintained in the future.

Could it be that the defence industry's automated vehicle developments come to benefit civilian applications, as has been the case with technologies such as GPS?

The concept of assuring a vehicle's behaviour throughout its lifecycle is a particularly relevant one to AI. While traditional vehicles are expected to perform consistently throughout their life, self-learning algorithms and fleet-wide software updates mean that an automated vehicle might handle the same scenario differently from one week to the next.

Another fundamental point to consider is whether the right things are being automated. The aim of the digital battlefield is to assist human decision makers to go beyond their current capabilities, whereas there's a view in some quarters that fully automated driving hasn't yet accomplished anything that a capable human operator can't already achieve. Should the automotive industry therefore look to other industries, such as aerospace, and focus on how AI can be used to augment the human driver's capability? For example, helping older drivers to keep driving safely, if their eyesight fades or their range of movement becomes restricted.

Inevitably, the main driver in military technology remains conflict. Recent conflicts have undoubtedly added urgency to the development of automated vehicles in defence. But could it be that the defence industry's automated vehicle developments come to benefit civilian applications, as has been the case with technologies such as GPS? The potential is certainly there.

Dr. Helen Monkhouse
Chief Engineer for Functional Safety at HORIBA MIRA

