Autonomy, Trust, and Transportation

Andrew R. Lacher
The MITRE Corporation, McLean Virginia 22102
alacher@mitre.org

Dr. Robert Grabowski
The MITRE Corporation, McLean Virginia 22102
rgrabowski@mitre.org

Dr. Stephen Cook
The MITRE Corporation, McLean Virginia 22102
scook@mitre.org

Abstract
Automation in transportation (rail, air, road, etc.) is becoming increasingly complex and interconnected. Ensuring that these sophisticated non-deterministic software systems can be trusted and remain resilient is a community concern. As technology evolves, systems are moving increasingly towards autonomy where the “machine” is intelligent: perceiving, deciding, learning, etc. often without human engagement. Our current mechanisms and policies for oversight and certification of these systems to ensure they operate robustly in safety-critical situations are not keeping pace with technology advancements. How is an autonomous system different than an advanced automatic system? How is trust different for an autonomous system? What are different perspectives on trust? Is it appropriate to apply the techniques used to establish trust in today’s transportation systems to establishing trust in an autonomous system? This paper will examine these questions and propose a framework for discussing autonomy assurance and trust in transportation applications. We will explore further with two examples: 1) the notion of a self-driving taxi-cab; and 2) the evolution of a two-pilot flight deck, to single-pilot operations.

Background
Often technology which originates or is perfected for military applications soon finds its way into a variety of commercial and civil applications. This is especially true of aviation. Unmanned vehicles have recently transformed U.S. military operations with remotely piloted aircraft and tele-operated robots becoming almost ubiquitous in the battle space. The U.S. Marine Corps has even deployed in Afghanistan an “autonomous” air cargo delivery aircraft for operational evaluation purposes (McLeary 2012). The U.S. Navy is in the process of procuring an “autonomous” air combat aircraft (Alkire 2010), while the U.S. Army is experimenting with leader-follow technologies for cargo logistics (Schoenherr 2009, Department of Defense 2013).

In the commercial world, many leading companies are exploring the application of autonomous systems to a variety transportation related applications. Most automotive manufacturers are routinely producing vehicles with varying levels of “driver-assist” technologies. Many major manufacturers have announced initiatives to explore even more sophisticated technologies moving towards driverless cars (Ford Motor Company 2013; Nissan Motor Company 2013; Toyota Motor Sales 2013). Google claims that they have driven their prototype driverless cars 500,000 miles without incident (Fisher 2013).

Due to the potentially dire consequences of failures in the transportation area, there are a number of processes in place that ensure that new technologies are safe. Regulators, manufacturers, and consumers themselves all play a role.

What is an Autonomous System?
It is often difficult to discuss an issue without a clear definition of the subject under discussion; however, the definition of autonomy or an autonomous system is not a simple matter. Autonomous systems decide for themselves what to do and when to do it (Fisher, Dennis, and Webster 2013). There is a range of functionality that is traditionally performed by people that could potentially be conducted through automation. At what point in this range a system becomes autonomous is difficult to define. Recently a Defense Science Board panel concluded that they could...
discuss autonomy without necessarily defining it (Defense Science Board 2012) and we will attempt to do the same.

**Characteristics of an Autonomous System**

There are some characteristics which are typically found in systems labeled at being autonomous. Note, for a system to be considered an autonomous system, not all of the characteristics listed need be present, just a preponderance. Some of these characteristics are shared by less sophisticated automation systems. As the automation increases in complexity the system begins to share more and more with an autonomous system. Table 1 contrasts the characteristics of an advanced automatic system with those of an autonomous system. When discussing an autonomous system, we will assume that the majority of these characteristics are present.

Autonomous systems in transportation will have varying degrees of sophistication as well as present test and evaluation challenges for developers, acquirers, and certifiers. Characteristics of autonomous systems that make them especially challenging are: 1) they are non-deterministic; and 2) employ adaptive algorithms. Traditional mechanisms for exhaustive testing will not work for a system that may make different decisions given the same input with all of them potentially correct. Autonomous systems perceive, judge, decide, and act. Determining how people can trust these systems for transportation is critical for enabling them to be used appropriately.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Advanced Automatic</th>
<th>Autonomous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reacts at cyber speed</td>
<td>usually</td>
<td>usually</td>
</tr>
<tr>
<td>Reduces tedious tasks</td>
<td>usually</td>
<td>usually</td>
</tr>
<tr>
<td>Augments human decision-makers</td>
<td>usually</td>
<td>usually</td>
</tr>
<tr>
<td>Proxy for human actions or decisions</td>
<td>usually</td>
<td>usually</td>
</tr>
<tr>
<td>Robust to incomplete missing data</td>
<td>usually</td>
<td>usually</td>
</tr>
<tr>
<td>Reacts to the environment</td>
<td>usually</td>
<td>usually</td>
</tr>
<tr>
<td>Exhibits emergent behavior</td>
<td>sometimes</td>
<td>usually</td>
</tr>
<tr>
<td>Adapts behavior to feedback - Learns</td>
<td>sometimes</td>
<td>usually</td>
</tr>
<tr>
<td>Responds differently to identical inputs</td>
<td>sometimes</td>
<td>usually</td>
</tr>
<tr>
<td>Addresses situations beyond the routine</td>
<td>rarely</td>
<td>usually</td>
</tr>
<tr>
<td>Reduces cognitive workload for humans</td>
<td>rarely</td>
<td>usually</td>
</tr>
<tr>
<td>Replaces human decision-makers</td>
<td>rarely</td>
<td>potentially</td>
</tr>
<tr>
<td>Robust to unanticipated situations</td>
<td>rarely</td>
<td>usually</td>
</tr>
<tr>
<td>Behavior is determined by the experience, rather than by design</td>
<td>usually</td>
<td>usually</td>
</tr>
<tr>
<td>Adapts behavior to significant environmental changes</td>
<td>usually</td>
<td>usually</td>
</tr>
<tr>
<td>Makes value judgments – weighted decisions</td>
<td>usually</td>
<td>usually</td>
</tr>
<tr>
<td>Makes mistakes in perception and judgment</td>
<td>potentially</td>
<td>usually</td>
</tr>
<tr>
<td>Introspective – Reviews performance to identify own mistakes</td>
<td>usually</td>
<td>usually</td>
</tr>
</tbody>
</table>

*Table 1. Contrasting the Characteristics of Automatic and Autonomous Systems*

**What is Trust?**

If an autonomous system is to operate in situations where the consequences of ineffective performance could cause physical harm to persons or property (e.g., unmanned aircraft, driverless car), humans will need to have a mechanism for both establishing and maintaining trust in the perception and judgment of these systems. People will need to be confident that the system will perceive the situation appropriately under all circumstances and make the correct judgment given the situation. **Trust is not a trait of the system; it is the status the system has in the mind of human beings based upon their perception and expectation of system performance.** Trust is a belief that something is expected to be reliable, good, and effective. Trust is based upon evidence and perception. Establishing and maintaining trust is not just an engineering challenge; it is a human factors challenge involving cultural,
organizational, sociological, interpersonal, psychological, and neurological perspectives (Lee and See 2004).

Trust is fundamentally a “mental state, a complex attitude of an agent x towards another agent y about the behavior/action a towards goal g.” Consequently only cognitive agents can trust (Castelfranchi 2000). Rempel, Homes, and Zanna studies the factors that lead to establishment of trust between humans in close relationships (Rempel, Homes, and Zanna 1985). They proposed that human-to-human trust relationships were based on predictability, dependability, and faith. Muir extended these principles to evaluating trust between humans and machines through experiments (Muir 1987). Lee and Moray conducted further experiments that suggested that “[...] the theoretical development of trust between humans presented by Rempel et al. (1985) also applies to trust between humans and machines” (Lee and Moray 1992). Lee and See further found that when people trust they are willing to put themselves or something they value at risk (Lee and See 2004).

In this paper, we will focus our discussion on how people establish trust in autonomous systems. As we increase the complexity of automation systems moving towards autonomy, the mechanisms for establishing human trust in these systems will rely even more heavily on the mechanisms used for establishing trust in people.

**Perspectives on Trust**

There are many different perspectives on trust depending upon a person’s role in interacting with an autonomous system. They will be asking themselves somewhat different questions when determining whether to put trust in an autonomous system because they are risking different things. At one end they risk their own lives; at the other they may be risking their reputation or personal finances. Using the metaphor resulting from the old story of a group of blind men and an elephant, some different perspectives are identified in Figure 1. Table 1 summarizes the roles and the questions that an individual might be asking in determining whether to trust a system. Most importantly, the categories of what might potentially be at risk are identified.

<table>
<thead>
<tr>
<th>Role/Questions</th>
<th>What is at Risk?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Researcher</strong> - What technology and methods are needed for establishing trust in autonomy and its safety?</td>
<td>Reputation</td>
</tr>
<tr>
<td><strong>Regulator</strong> - Can I confirm the performance sufficiently? Is it safe?</td>
<td>Reputation, Job security, Public trust</td>
</tr>
<tr>
<td><strong>Creator</strong> - How should I design/architecture the system so that I can demonstrate its trustworthiness? What liability do I have if the system fails? Is it safe?</td>
<td>Reputation, Job security, Employer’s finances</td>
</tr>
<tr>
<td><strong>Insurer</strong> - How do I go about establishing underwriting rates? Is it safe?</td>
<td>Job security, Employer’s finances</td>
</tr>
<tr>
<td><strong>Community</strong> - Do I want this system operating in my neighborhood? Can I count on it? Is it safe?</td>
<td>Personal safety, Personal property/finances</td>
</tr>
<tr>
<td><strong>Acquirer</strong> - How do I know that the system meets my requirements and can be relied upon? Is it safe?</td>
<td>Job security, Employer’s finances, Mission effectiveness</td>
</tr>
<tr>
<td><strong>Commander/Supervisor</strong> - What changes operationally? Can I rely upon this technology in operations? Will I be responsible for incidents that may result from operations? Is it safe?</td>
<td>Job security, Mission effectiveness, Personal safety, Personal finances</td>
</tr>
<tr>
<td><strong>Operator</strong> - What can I rely upon this system to do? Will I be responsible for incidents that may be result from operations? Is it safe?</td>
<td>Job security, Mission effectiveness, Personal safety, Personal finances</td>
</tr>
<tr>
<td><strong>Patron</strong> - Should I use it? Is it safe?</td>
<td>Personal safety, Personal finances, Personal property</td>
</tr>
</tbody>
</table>

**Figure 1. Perspective of Trust Vary by Role**

Trust is not narrowly limited to direct interaction between two agents; where one agent trusts the other to behave in its best interest. Trust can be gained by the introduction of a third party that takes some investment in the relationship between the two parties. This third party may provide oversight, regulation or enforcement of a contract (social, legal, or both) between the two parties. This complex relationship between two agents in the presence of a third is defined as three party trust (Castelfranchi and Falcone 2000).
Sometimes the fact that a regulator or some authority has expressed trust in a system is sufficient for other individuals to gain trust of these systems (Schilke and Cook 2013). For example, passengers do not typically scrutinize an aircraft manufacturer’s engineering processes to determine whether a particular aircraft is safe. Rather they rely upon the fact that the FAA has certified the aircraft for flight. Similarly, passengers do not interrogate pilots and crew but rather gain confidence knowing the infrastructure of training that pilots must go through before being allowed to fly an aircraft. Finally, confidence is also bolstered by understanding the financial litigation cost of a company failing to take appropriate safety measures. Each of these in itself is not sufficient to build trust but can greatly enhance it when present.

**Multi-Party Trust**

We extend the notion of three party trust to our perspective model to introduce the concept of multi-party trust. The means by which trust is established from each perspective differs significantly by that party’s role. As an example, establishing trust from the perspective of a regulator (e.g., the Federal Aviation Administration) is very different from ensuring trust from the perspective of an operator (e.g., the pilot) or the community (e.g., population on the ground). By decomposing the different roles and perspectives, we gain an actionable means for building trust across a community. We are reminded of the African proverb ‘it takes a village…’; for developing trust in autonomous systems, ‘it takes a village to raise and trust an autonomous system’ because we require diverse elements of the community to establish trust.

**Certification**

Certification is a formal means by which a regulator confirms expected performance of a system built by a manufacturer. This confirmation is based upon a review of evidence resulting from formal engineering processes that are well established. By certifying a system, the regulator is essentially saying that the system can be trusted to perform as designed. Two examples of certification of transportation systems are examined below – civil aviation and submarines – and implications of autonomous systems on these certification processes are considered.

**Example – Civil Aviation**

The U.S. civil aviation industry depends upon the Federal Aviation Administration to establish the airworthiness of aircraft before they can commercially operate. While airworthiness certification is only one aspect of operational safety, it has led to civil aviation being regarded as the least dangerous form of transportation (Jacobs 2013). Downer found that as the complexity of aircraft has increased, the FAA has relied more heavily on the experts who build the system to attest to the aircraft’s airworthiness (Downer 2010). Downer states that “The FAA type-certification process is important, but aircraft are complex and inscrutable and so auditing them must lean heavily on the tacit knowledge of the engineers who build them.” As systems move toward complexity and autonomy, this trust relationship between the regulator and the manufacturer will likely grow in importance.

**Example – US Navy Submarines**

The US Navy implements safety for its submarine fleet through the SUBSAFE program. The SUBSAFE program is considered very effective – since it was implemented in 1963 the U.S. Navy’s submarine loss rate has gone from one per every three years to zero in the years since (Leveson, 2011). An important element for the certification portion of the SUBSAFE program is its reliance on Objective Quality Evidence – verifiable, testable observations, and measurements used to form the certification basis. Recent Navy Congressional testimony, revealed that for the SUBSAFE program “Without objective quality evidence there is no basis for certification, no matter who did the work or how well it was done” (Sullivan 2003). As we stated earlier in this paper, autonomous systems are characterized by non-determinism and employment of adaptive algorithms. To employ effective certification programs that rely on objective quality evidence, it will be critical to develop methods and processes to produce measurable, verifiable test results for autonomous systems.

**Framework**

In addition to considering the varying perspective of trust in autonomy, we found it useful to develop a framework to provide a convenient vocabulary for discussing various aspects of trust. The framework appears in Figure 2.
Fundamentally we structured our thinking around three major factors: 1) the autonomous system; 2) the people who interact with it; and 3) the environment in which it is intended to operate. We will discuss each factor below.

People
We will start with people since they are the most important part. Autonomous systems are designed and constructed by people; they are regulated, managed, and operated by people; ultimately for the benefit of people.

As we discussed above, people’s perspectives of the autonomous system will vary based upon their role. One thing in common is that people’s perception will be based upon evidence that they can observe; evidence that is presented to them; and their own basis which includes not only their experience but a cultural perspective. How an individual perceives system behavior will be greatly influenced by their culture (Bailey, Gurak, and Konstan 2001). Culture has many dimensions to include: ethic, religious, national upbringing, professional affiliation, and age. Let’s consider two examples: A teenager who grew up using tablets, smartphones, and the internet is much more likely to have confidence in the ability of a system which is capable of automatically parallel-parking a car than a gentleman in his 60s who still changes his own oil. Also consider a situation with more dire consequences. Apparently, the National Transportation Safety Board is investigating whether Korean culture of respect may have been a factor in the recent ASIANA crash in San Francisco. Apparently, the First Officer recognized a problem with the automation’s ability to maintain the proper glide slope but out of difference to (or perhaps because of trust in) authority did not press the issue (NTSB 2013).

System
The system was designed and constructed to have a specific level of competency. The real competency or trustworthiness exists whether it can be accurately observed, measured, or assessed. It is a trait of the system. This real competency is going to be dependent upon the system’s architecture and the quality of the execution of the development activities. Competency is an engineering trait. There are many ways by which a systems competency can be estimated. For most of today’s systems, intended functions can be tested and assessed. People’s confidence in an automation systems functionality will be only partially dependent upon the results of these tests. Confidence is a human trait. If in testing, known inputs produce expected outputs/behaviors, people may be confident that in operation the system will function as intended. Basically does the system do what the human intended, when he intended it to do it (Fisher, Dennis, and Webster 2013)? Exhaustive testing is just one method to determine a systems competency. A less formal method might include simply observing how well the system works in operations.

Most systems involve some degree of human interaction or collaboration. To be effective, even highly automated and autonomous systems will be interacting with humans. Often people’s confidence in a system can be determined not just by the system doing what it should and when it should, but it needs to be doing it for the right reasons (Fisher, Dennis, and Webster 2013). Thus the nature of the human-machine interaction (i.e., the collaboration) needs to be established so that the human operator understands why the system is functioning the way that it is. This will help determine confidence in that system’s ability.

“There is nothing worse than a so-called smart machine that can’t tell you what it’s doing, why it’s doing something, or when it will be finished.” (Bradshaw, et al 2013)

Environment
The environment a system is intended to operate will have significant influence on the system. First, environmental circumstances will establish the behavior of the autonomous system. The autonomous systems’ perception of the environment will be determined by the content of the data received from its sensor inputs. Correctly perceiving the circumstance is an important part of an autonomous systems ability to respond correctly.

The system may have bounds or constraints on its behavior as part of the system design to limit negative effects or consequences on people of poor system performance. As confidence in the system grows, these constraints could be relaxed. An approach to consider is incrementally integrating an autonomous system by initially constraining system behavior as a means to limit consequences of system failures until confidence grows. Effectively we would be matching authority with the level of demonstrated trustworthiness of the system. This is very similar to how we manage Extended-range Twin Operations (ETOPS). As confidence in the reliability of an engine grows we give the operator increasing authority to transit further from suitable diversion airports (FAA 2008). Using an approach that matches autonomous system authority with the level of earned trust follows societal norms as well as practices established in the transportation domain.

A Taxi Ride
As a thought experiment to exercise the framework and illustrate the varying perspective on trust in systems, we
would like to consider two examples. The first is a ride in a taxi cab.

**A Ride in a Taxi Today**

Consider the following scenario:

You are in a large US city and you need to flag down a taxi cab for a ride to the airport during rush hour. You stand on the street corner and wave. A car moves to the right to pick you up, you open the door, enter, and tell the driver your destination.

There are a number of factors at work that make you trust a ride from a total stranger in an unfamiliar city. Let’s consider them. You actually made a number of subconsciously and conscious assessments of the taxi-driver, the taxi, and entire situation before you entered the cab. Most likely you did not even start to flag down a cab until you saw a car that was painted like a cab. Perhaps you only waved at the newer looking cabs or only the cabs from a specific company. Before getting into the cab you saw that it was painted like a cab, had clearly identifiable corporate markings, had a license plate, and was in reasonable condition. Perhaps you also considered the size of the car, the age of the car, and the appearance. You also looked over the driver. Was he/she dressed okay? Not too old and not too young? Did he get out of the car to help with bags?

By entering that cab, you were putting trust not only in the cab driver but in an entire system that led to the cab being available. You were essentially also trusting that the particular taxi company hired good drivers and maintained their vehicles. You trusted the brand name manufacturer of the taxi, the builders of the road, the state that licensed the driver, the authority that gave him a hack license, the authority which gave an operator license to the cab company, etc. (Carlson et al 2014)

You were also trusting that the person driving the cab did not want to die, did not want the cost of an accident, and did not want to lose his license (and thus his livelihood) by breaking laws (e.g., speeding, drinking while driving, running lights/stop signs, etc.). You were fairly certain that there were police monitoring driver behavior, and that the taxi company would fire bad drivers. You trusted the driver, the taxi company, the hack commission, the city, the state, the manufacturer… the system… because you understand their motives and their relationships. This is an example of multi-party trust.

You also knew that there were constraints on the consequences of failure. The distance and thus exposure was not that great, you would not be going all that fast, you buckled your seat belt, the car has air bags, and other safety features. If the driver tries to kidnap you, you have a cell phone to make an emergency call. Some people avoid flagging down random cabs on the street corner and will only use cabs that have been dispatched by the company based upon a phone call, or were found at hotel taxi stands. This provides another layer of oversight and another party that can be trusted.

How would “trust” in the ride change if the car stopping was 25 years old, had dings/dents, was missing its bumper, had bald tires, and a hand-stenciled logo. What if the driver did not even look 18 years old, was lounging in the seat, and slurred his speech. What if the cab careened across three lanes of traffic to stop and pick you up? What if you were in a foreign country? Would you still get in a car for a ride with a total stranger?

Would you accept a ride from a cab you trusted less if the environmental circumstances were different? If it was raining, or you were late for your flight, or it was nighttime and you had not seen a cab for a long time?

**Driverless Car for Hire**

Let’s now consider a moment in the future with the same scenario; the only difference is that the car stopping to pick you up is a driverless taxi cab.

You can still judge the appearance of the vehicle but there is no driver to judge. You will need to trust that the manufacturer of the driverless car built an effective system because its investors never would have wanted to sell a system that could make them liable for significant damages. Perhaps you would have confidence in only certain brand name driverless car manufacturers (Carlson et al 2014). You have to trust that the taxi company never would have purchased a driverless cab unless it functioned as it should. You would need to trust that the hack commission, the city, and the state would never have permitted these driverless cars on the road unless they worked. You have to trust that all of these parties monitor the performance of these vehicles and if they demonstrated operational issues, they would have been removed from service.

Maybe there would be a readily accessible “stop button,” like an emergency brake on a train that would enable you as a passenger to gracefully stop the taxi if something was not right. This would serve as a means for you as the patron to directly constrain behavior. You don’t have that on a taxi today and perhaps sometimes wish you did.

Right now no regulatory authority exist that would certify the autonomous system operating the driverless car. One of the major governors on deployment of the technology is desire for designers (e.g., auto manufacturers) and acquirers (e.g., taxi companies) to avoid liability claims resulting from premature deployment of a technology that might be faulty.
Single Pilot Operations

As another thought experiment let us contrast flight operations today with a transition towards a single pilot operation.

Commercial Flights Today

In accordance with 14 CFR Part 25 – Airworthiness Standards: Transport Category Airplanes, a manufacturer must establish the minimum number of flight crew members (i.e., pilots) required for safe operation (FAA 1965a). Many factors are considered including workload, accessibility of controls, the kind of operation, process complexity, use of automation, failure conditions, and emergency procedures (FAA 1965b). One factor that is important in determining the minimal flight crew required is whether there will be the need to perform multiple functions at the same time (FAA 1996). Most commercial operations follow the operational rules defined by 14 CFR Part 121 – Operating Requirements Domestic, Flag, and Supplemental Operations which makes it clear that “the minimum pilot crew is two pilots” (FAA 1996). One pilot is designated pilot in command (FAA 1996) who has “final authority and responsibility for the operation and safety of the flight.” (FAA 1962). Operational duty limitations may require additional flight crew members depending upon the duration of the flight enabling other crew members to rest. For the purpose of this discussion we will just focus on the number of crew members to operate the aircraft.

For most flight decks, there is a concept of a pilot flying the aircraft and the pilot-monitoring also known as the pilot-not-flying. Thus, one pilot has their hands on the controls, while the other is assisting by monitoring what the first pilot is doing. The pilot in command determines who is performing which role. The exact procedures for the pilot-flying and pilot-not-flying are defined in the Flight Operations Manual developed by the operator (e.g., airline) and approved by the FAA in accordance with the aircraft operations manual developed by the manufacturer.

Today there is considerable automation certified by the FAA and available on the flight deck. There has been much written about human factors challenges associated with the trustworthiness of this automation and whether the human-machine interface is effective including identification of problems with mode confusion (Flight Deck Automation WG 2013, Traufetter 2012).

In general, pilots, airlines, and regulatory authorities have established a level of confidence in the existing system automation based upon initial certification activities, demonstrated performance, and collaboration with the system during operations. There may be competency limits in today’s systems but given the constraints and limitations on their use (usually through training and procedures) the consequences of their failures are generally within acceptable levels of risk.

Transition to Single Pilot Operations

To improve operational efficiency and safety, research to transition from a two-pilot flight deck to single pilot operations is underway (Deutch and Pew 2005, ACROSS 2013, Learmount 2103). Transitioning to single pilot operations of commercial aircraft which achieves the level of safety and operating efficiency of today’s two person flight decks presents significant challenges. As we design single pilot operations, we will need to keep in mind that today’s two-person flight deck with well-established roles of pilot in command, pilot-flying, and pilot-not-flying are still vulnerable to errors that lead to incidents and accidents (Deutch and Pew 2005).

If we are to move towards operations where a single pilot can be trusted to safely operate an aircraft in commercial operations a number of things will have to happen.

- The regulator would need to change existing regulations based on clear safety evidence and a formal risk assessment.
- Manufacturers will need to develop system solutions that would enable a single pilot to operate the aircraft perhaps by automating many of the pilot-not-flying functions and constraining the actions taken by the automation to limit consequences of failures; designing the system so that the pilot only has one thing to do at a time.
- Operators will need to develop operational procedures and training to prepare pilots for single pilot operations.
- Manufacturers and operators will need to develop the data necessary to demonstrate to themselves and the regulator that they can have confidence in the system’s competency.
- Researchers will need to develop methods and procedures for measuring the competence and safety of the autonomous system.
- The Regulator will need to certify the systems, procedures, and training associated with the single pilot operations; expressing their confidence in the competency of the system and its ability to perform its intended functions.
- Insurers will have to underwrite the operation.
- Finally, the general public and other pilots will need to have confidence that these systems will perform as expected. They will need to trust third parties like the regulator, developers, and operators and the trust they have put in the system.
Observations and Conclusion

Autonomous systems have characteristics such as being non-deterministic and employing adaptive algorithms. We explored some of the unique challenges that autonomous transportation systems present with regard to traditional certification approaches. However, trust is not a trait of the system; it is the status the system has in the mind of human beings based upon their perception and expectation of system performance (i.e. the systems competency). Trust is essentially the level of confidence people have in the system based on their observations of the system operations and other data produced as evidence of its competency.

There are multiple perspectives of trust which are determined by the role and what individuals have at risk. We discussed the concept of multi-party trust and how it can be extended to a framework illustrating the relationships between disparate roles. We presented two thought experiments showing that building and maintaining trust in the perception and judgment of increasingly autonomous systems will be a challenge for the transportation community. In order to establish trust in autonomous systems, a clear understanding of the many ways in which people, the system, and the environment interrelate will be essential.

Disclaimer

The views, opinions, and/or findings contained in this presentation are those of author(s) and The MITRE Corporation and should not be construed as an official Government position, policy, or decision, unless designated by other documentation.

APPROVED FOR PUBLIC RELEASE - Distribution Unlimited – Case Number: 14-0001.

References

Toyota Motor Sales 2013. Toyota to Launch Advanced Driving Support System Using Automated Driving Technologies in Mid-2010s, Press Release.