AV Combinator

Autonomous Vehicle Modular Safety Suite

A New Approach to Autonomous Vehicle Safety

White Paper

April 2021

Autonomous Vehicle Modular Safety Suite: A New Approach to Autonomous Vehicle Safety

White Paper

April 2021

Authors:

Michael Nehmad
Jenn Hu
Richard Sun
Troy Lawrence
Julius Niehaus

Contents

Executive Summary

The biggest gating factor to widescale adoption of Autonomous Vehicle (AV) technology is safety. Increasing public confidence in AV safety will accelerate the adoption of AVs. In this white paper, we attempt to reframe vehicle safety through a large-systems approach – one that is meant to aggregate and integrate all possible approaches to AV safety into a single, holistic framework, such that the whole is greater than the sum of its parts. Additionally, this framework will create a clear, concise, and common safety language for all stakeholders in the AV ecosystem such that all parties can have an active role in shaping the future standards of AV safety through a process of natural convergence.

Why a New Approach to AV Safety is Necessary

AVs present both tremendous opportunity and potential risk to our way of life. On one hand, AVs may save tens of thousands of lives each year from automobile, return countless hours of wasted time to Americans, and decrease dangerous physical interaction in a global pandemic. On the other hand, there is risk in turning over control of a safety-critical task to autonomous systems as the decisions they make will impact the lives of not just its users, but that of all road-goers.

In order to bring about the benefits and mitigate the risks of AVs most effectively and expeditiously, a more complete picture of AV safety is needed. Traditional vehicle safety has focused mostly on passive safety features that protect occupants in the event of a crash, such as the vehicle's body structure, in addition to airbags and seatbelts. While protective features remain critical, AVs present a new paradigm of preventative, active safety features that may enable a majority of crashes to be avoided altogether.

However, given the nascency of autonomous technology, it has been difficult to accurately showcase and assess the level of safety provided by these new preventative features. Over the years, the California Department of Motor Vehicles autonomous vehicle safety reporting and licensing program has received considerable pushback from the AV industry regarding its use of disengagement metrics as a proxy for technology maturity. The industry has also yet to convene on a single approach; AV companies continue to disagree amongst themselves regarding which safety metrics should be used to indicate system safety. Similarly, in areas where AV testing is being conducted, AV companies and local governments have communicated to the public that AVs are safe, even though researchers agree that reliable data of crashes and near misses are "unable to provide information about safety." Given this state of confusion, the public remains rightfully skeptical of self-driving vehicles.

Many of these existing solutions share one common feature: they are small-systems approaches to safety, in other words they evaluate AV safety through a singular lens, method or metric. Narrowly defining AV safety can simplify regulatory decisions, but at the same time, risks creating blind spots, by neglecting the nascent, complex, and non-deterministic nature of AV technology. What is needed, yet missing, is a large-system approach to AV safety.

The objective of this white paper is to conceptually develop this large-systems framework for the improved understanding and assessment of AV safety. In doing so, it aims to: (1) level-set all parties on the possible ways protective and preventive AV safety can be showcased or evaluated; (2) create a common safety language that is accessible and comprehensible by all parties, and (3) educate all parties as to how they can best leverage this framework for their particular objectives. Where the small-systems approach struggles to handle the interconnected complexities of this new technology through unilateral resolutions, the large-systems approach thrives by accepting these complexities, and providing all parties with the right tools and information to enable them to converge towards the best resolution collectively over time.

Introducing the AV Modular Safety Suite: A large-systems approach to AV safety

The MSS seeks to collate all the possible approaches one could take to define a particular aspect of safety as it relates to an AV. Here, each small-systems solution is effectively a “module”. Using this approach, the user can choose from a pool of existing modules to develop their own proprietary “suite” based on their priorities, existing policies, and risk tolerance. By providing the information and tools to create quickly customizable suites, the MSS encourages experimentation which can accelerate the convergence towards a set of global safety standards, or at the very least, a collection of industry best practices.

This proposal presents a major change to past approaches. Instead of viewing existing safety solutions and standards in silos, it aims to aggregate these stand-alone variables into one function with an output that can be used to holistically assess the safety of AVs. In addition, the MSS will also outline pathways to evaluating future components of the safety function that do not yet exist, but are helpful in comparing the safety performance of AVs against that of a human driver and against a user's defined standard of safety. Such findings can then be employed to empirically prove the safety benefits AVs will bring to broader society.

Guiding Principles of the MSS

The MSS will be technology neutral: The MSS will only consist of modules that do not specify the use of a certain technology in creating the automated driving system. This aligns with the spirit of the current US DoT policy that aims to provide an open regulatory environment to encourage innovation as a means of achieving safety.

The MSS will be politically agnostic: The MSS intends to be flexible so that it adds value regardless of any administration that is in office. At its worst, the MSS provides a pathway that enables all parties to share a common safety language and improve public understanding of a nascent and complex technology. At its best, the MSS equips stakeholders with the data and tools needed to craft effective regulation.

The MSS should be continually updated to be as comprehensive as possible: Given the nascency of the industry, it is likely that new approaches and solutions for AV safety will be developed in the future. In these cases, the MSS is designed such that it can be updated to include them.

The MSS will be open and transparent: All information and data published through the MSS will be open and free to all to facilitate public education on the multiple facets of AV technology and safety. That is not to say that the MSS is responsible for enforcing information sharing requirements on AV companies, should they be enacted into legislation by the local or federal governments. Rather, it means that any published material should be made as widely available to the public.

The MSS will be centrally located: All information and data published through and about MSS should be accessible via a centralized platform. This creates a one-stop-shop for anyone trying to learn about AV safety. Additionally, it will make tracking statistics, such as how many AV companies are utilizing a particular module of the MSS, much easier.

The MSS should be comprehensible by the general public: All data and information stored on the centralized platform should be presented in a way that is simple and comprehensible by all, as the safety-critical nature of the technology can affect the lives of all road-goers, not just AV service users.

The MSS aims to develop industry-wide best practices via data-driven convergence: Rather than a central body winning favor through traditional means like lobbying, the MSS enables all stakeholders in the AV ecosystem to play an active role in determining best practices and future AV safety standards by facilitating and showcasing data-driven coordination across the entire ecosystem.

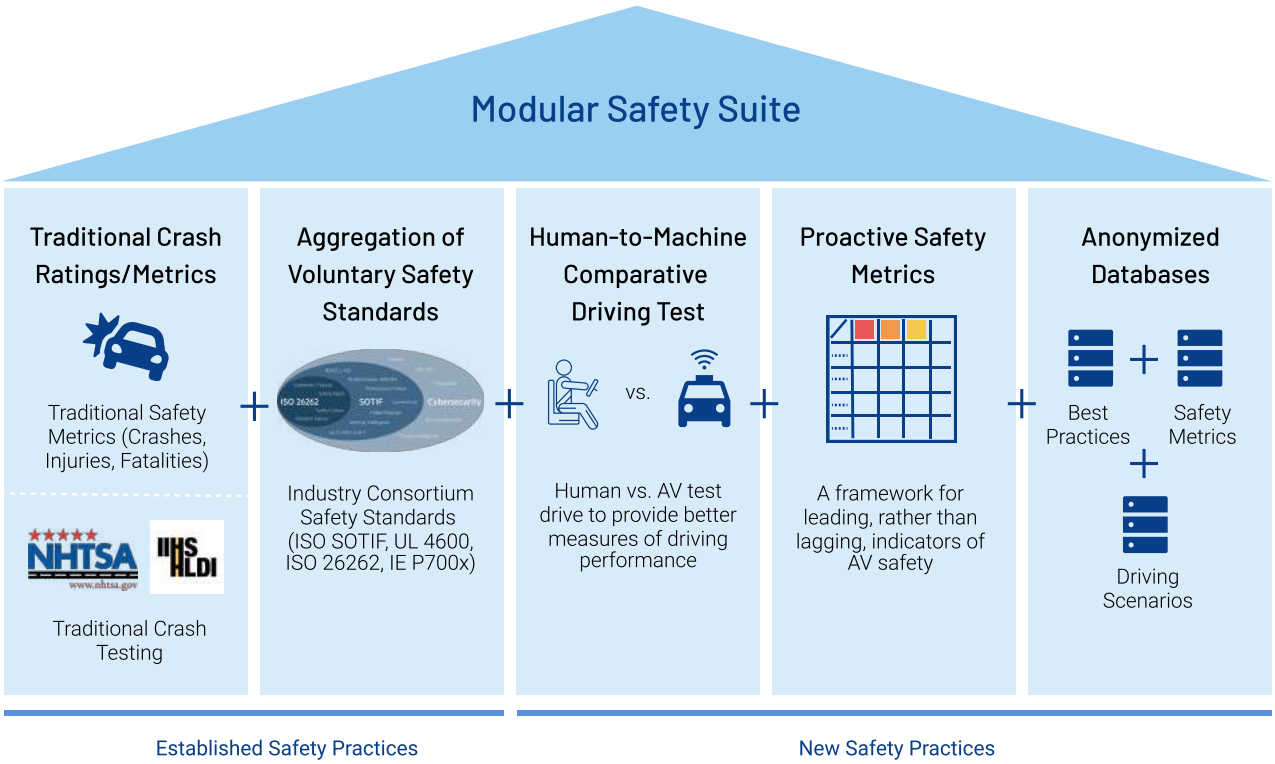
Overviewing the Modules in the MSS

The MSS is currently comprised of six modules, each of which is covered in detail in their respective sections:

1. Traditional Safety Metrics
2. Traditional Crash Testing
3. Aggregation of Voluntary Safety Standards,
4. Proactive Safety Metrics
5. Human-to-Machine Comparative Driving Test
6. Anonymized Databases of Best Practices, Safety Metrics, and Driving Scenarios

The six modules in the MSS are split between lagging measures that track incidents and outcomes, as well as leading measures that predict future outcomes. Leading measures are proxies for lagging measures, which while less concrete, provide foresight to the technologies' performance prior to deployment. By encompassing both types of measures, the MSS intends to produce an output that gives a holistic and comprehensive view on AV safety (Figure 1).

Figure 1
Overview of the Modular Safety Suite



1. Traditional Safety Metrics

Traditional safety metrics refer to the recording of vehicle crashes, injuries, and fatalities. These metrics have been the benchmark for measuring safety in the automobile industry.

In the context of AVs, traditional safety metrics remain relevant as they not only allow for the comparison of real safety outcomes of autonomous driving to human driving, they also provide data which can be used to validate the accuracy of the other MSS modules in their assessment of AV safety.

2. Traditional Crash Testing

Traditional Crash Testing involves testing the vehicle's physical structure to ensure that the vehicle is designed in a manner that sufficiently protects its occupants and other road-goers. This module will build on two existing crash systems in particular, the 5-Star Safety Ratings Program created by NHTSA's New Car Assessment Program (NCAP), and the Insurance Institute for Highway Safety's (IIHS) Vehicle Rating System.

This module also serves to remind the general public that AVs will not be impervious to crashes despite their improved safety potential.

3. Aggregation of Voluntary AV Safety Standards

This module aims to consolidate existing and future AV safety standards put forth by various Standard Setting Organizations (SSOs), while tracking the utilization of each standard across the AV industry.

The utilization of each standard by AV players can then be mapped to the safety metrics contained in other modules. This, in turn, can help SSOs identify the strengths and weaknesses of their approach such that they continue to improve their standards through an immediate feedback loop. Additionally, it will facilitate the convergence towards a set of best practices for governing agencies to consider for the formulation of policy.

4. Proactive Safety Metrics

Proactive Safety Metrics aim to highlight the positive safety attributes of AVs, instead of only focusing on the potential of AVs to minimize the negative attributes of human-driven vehicles. Specifically, this module proposes a framework for the creation of proactive safety metrics, which are metrics that can be accurately gathered and assessed without removing the safety driver to measure the AV's abilities in perception, prediction, planning, and execution.

This framework would allow stakeholders in the ecosystem to better understand and track the advancement of AV technology relative to human drivers, without having to remove the safety driver from AVs on public roads.

5. Comparative Driving Test

A driving test that puts a human driver behind the wheel of an AV to drive through an area that is within the AV's Operational Design Domain (ODD), while the autonomous system runs in the background.

This test assures that factors such as conditions and routes are kept constant such that the safety of the choices made by the human and the autonomous system on the road can be fairly compared.

6. Anonymized Databases of Best Practices, Safety Metrics, and Driving Scenarios

This module proposes to build anonymized databases of best practices, safety metrics, and driving scenarios that can be accessible to regulatory bodies and the general public.

The voluntary and anonymized nature of the databases ensures that AV companies are able to protect themselves, while creating an incentive structure that promotes information sharing to enable all parties to benefit from more data. This data could be used to train and validate AV technology on a wider, more comprehensive dataset, while supporting the formation of data-driven safety standards, ultimately leading to the safer deployment of AVs.

How to Read this White Paper

Each of the following sections of this white paper covers a single module in the MSS. It is recommended that readers review the overview of each module before deciding which individual modules to pursue.

Every section subsequently dives into the module's purpose, followed by its detailed description, the value it brings as a stand-alone module and finally, the role it plays within the large-systems framework.

Autonomous Vehicle Modular Safety Suite: A New Approach to AV Safety

1 Traditional Safety Metrics

Overview

Traditional safety metrics - the recording of vehicle crashes, injuries, and fatalities - have formed the benchmark for the measuring of safety in the automobile industry for decades. In the context of AVs, traditional safety metrics allow us to compare the safety outcomes of autonomous driving to that of human driving, while also providing data that can be used to validate the accuracy of the other MSS modules in evaluating the safety of AVs. Instead of developing new methodologies, federal and local governments should build on existing systems, specifically the Crash Report Sampling System (CRSS), the Fatality Analysis Reporting System (FARS), and the State Data System (SDS), to collect data on AV crashes and fatalities.

The Strengths and Limitations of Traditional Safety Metrics

The purported safety benefits of AVs are often referenced in the context of traditional safety metrics, which are reactive, lagging measures of vehicle incidents that involve actual harm; namely crashes, injuries and fatalities. Promising fewer crashes, injuries, and fatalities is a simple way for AV companies to communicate the value of AV technology, given the fact that these metrics have been used understand road safety for decades. Additionally, given the traceable, observable nature of these statistics, an AV company can claim that their vehicles are safer than human-driven vehicles by comparing the number of crashes, injuries, and fatalities an AV has incurred on a per-mile basis to those incurred by an average human driver, once a statistically significant amount of road data has been gathered.

However, despite the benefits of this straightforward approach, traditional safety metrics come with one critical limitation, which is that gathering such data poses an inherent, non-trivial amount of risk to all road-goers. On one hand, AVs would need to operate on public roads without a safety driver to gather the data (crashes, injuries, fatalities on a per-mile basis) required to evaluate their safety performance. On the other hand, allowing such AVs to operate on shared roads would be unacceptable as these AVs have not yet been determined to be safe, given that the necessary data has not yet been collected.

In fact, the premise of the MSS relies on the fact that traditional safety metrics alone are not sufficient to evaluate the safety of AVs in a manner that is safe for all road-goers. This is not to say that traditional safety metrics have no place in the AV ecosystem. To the contrary, these are some of the most valuable metrics that can be employed because of their empirical nature. The key to optimizing the value of traditional safety metrics, as it relates to AVs, is to include them as only one facet of a broader, more holistic AV safety framework, ideally one that includes other inputs that can compensate for the gaps that are inherent in these traditional safety metrics.

Integrating Traditional Safety Metrics into the MSS

The large-systems approach underlying the MSS enables complementary support across the various modules. While traditional safety metrics are direct indicators of safety, other modules in the MSS can only serve as estimates or proxies for safety. On the other hand, while traditional safety metrics require a safety driver to be removed from the AV to be accurate and valuable, other modules in the MSS can help to evaluate safety without removing the safety driver from AVs on public roads. By adopting a large-systems approach, relevant players can selectively utilize different modules in ways that best leverage each module's strengths, while mitigating their weaknesses.

For example, an AV company might choose to first gather safety data via the proxies outlined in Modules 3 to 6. The AV company can then choose to remove the safety driver to implement this module when the data from the other modules has instilled enough public confidence in the autonomous system.

Additionally, observable traditional safety metrics can act as a source of feedback for other modules to improve the accuracy of their estimative power over time. This is because the large-systems approach enables the establishing of relationships, if they exist, between the data collected across the different modules in the MSS.

As an example, once an AV company has gathered data on both traditional safety metrics as well as the safety estimates from other modules within the AV's Operational Design Domain (ODD), those datasets can be compared to identify the safety estimates that most strongly correlate with the real-world data gathered from the traditional approach. This will then endow even greater confidence in the estimative capabilities of those proxy metrics as the AV company scales to other cities. Through this, it will become clearer over time which modules in the MSS are more indicative of AV safety than others, accelerating the convergence towards a set of industry best practices for the evaluation of AV safety.

NHTSA's National Crash Data Systems

This white paper will draw on NHTSA's National Crash Data Systems to give a concrete example of how existing reporting systems can be modified such that they can be applied to AV safety.

At the present moment, NHTSA's crash data collection program consists of six different data collection systems. Out of the six, the MSS finds three to be particularly applicable to the scope of AV safety: The Crash Report Sampling System, the Fatality Analysis Reporting System, and the State Data System.

Crash Report Sampling System (CRSS)

The CRSS is a sample of police-reported crashes involving all types of motor vehicles, pedestrians, and cyclists, ranging from property-damage-only crashes to those that result in fatalities. It obtains its data from a nationally representative probabilistic sample, selected from the estimated 5 to 6 million police-reported crashes that occur annually. It should also be noted that by focusing attention on police-reported crashes, the CRSS concentrates on crashes that are of greatest concern to the AV ecosystem and the general public.

Given those factors, the CRSS is the most appropriate measure in NHTSA's toolkit for estimating the overall crash picture in the US every year, and its methodology should be applied to gain a similar understanding of AV safety.

Fatality Analysis Reporting System (FARS)

The FARS contains data on a census of fatal traffic crashes across the US. To be included in FARS, a crash must involve a motor vehicle traveling on a traffic way open to the public, and must result in the death of a vehicle occupant or a nonoccupant within 30 days of the crash.

FARS can be applied to AV safety to highlight, more starkly, the differences in safety outcomes between human and autonomous driving. This would be particularly useful for educating the public about the safety benefits for AVs, to build public buy-in for the widespread testing and deployment of AVs on public roads.

State Data System (SDS)

The SDS contains computer data files coded from police accident reports from the thirty-two states that are currently participating in the SDS. As regional specificities are lost in national data samples, expanding the SDS to include AV safety data is particularly useful in giving a regional overview of how AVs compare to each other and to human drivers. This would be more useful to state governments in the formulation of state-level regulations, as compared to data that is sampled on a national level.

Collecting Crash Data for Autonomous Vehicles

When collecting data surrounding AV crashes and fatalities, it is important to ensure that enough information is recorded to determine whether the crash or fatality was indeed caused by a failure in the autonomous system.

This is important especially when the data is being used to compare the performance of autonomous driving against that of human driving as confounding factors such as rescue facilities, road design, and reckless pedestrians, might make for unfair comparisons. The "Crash" and "Post-Crash" rows of the Haddon Matrix (Figure 2) provide a non-exhaustive list of factors that should be recorded and taken into account when making such comparisons. The same can be said for when the data is being collected to validate the assessment quality of other modules in the MSS that look at the safety performance of autonomous driving.

Figure 2
Haddon Matrix

Phase	Human Factors	Vehicle and Equipment Factors	Environmental Factors
Pre-crash	• Information • Attitudes • Impairment • Police enforcement	• Roadworthiness • Lighting • Braking • Handling • Speed management	• Road design • Road layout • Speed limits • Pedestrian facilities
Crash	• Use of restraints • Impairment	• Occupant restraints • Other safety devices • Crash-protective design	• Crash-protective roadside objects
Post-Crash	• First-aid skills • Access to medics	• Ease of access • Fire risk	• Rescue facilities • Congestion

Autonomous Vehicle Modular Safety Suite: A New Approach to AV Safety

2 Vehicle Crash Testing

Overview

Vehicle crash testing addresses another the physical, protective aspect of safety. The physical structure of the AV can play a role in either protecting or inflicting more damage on its occupants and other road-goers in the event of a crash. This module will look at two crash systems in particular, the NHTSA 5-Star Rating System and the Insurance Institute for Highway Safety's Vehicle Rating System, to give examples of how existing crash systems work and the way in which they can provide a comprehensive picture of a vehicle's ability to protect its occupants during a crash.

Why AVs Still Need Crash Testing

Most modules in the MSS focus on evaluating the safety performance of the autonomous system as a measure of preventive safety. In other words, they evaluate the AV's ability to avoid crashes, as well as other traffic accidents. However, even if autonomous systems were to function perfectly from launch, which is an unlikely occurrence in and of itself, crashes involving AVs would still be inevitable as these AVs would need to share the road with traditional, unpredictable human-driven vehicles for the foreseeable future. In this light, to create a truly holistic view of an AV's safety performance, during- and post-crash safety performance need to be accounted for.

Fortunately, there already exist reliable and proven measures of protective safety performance in the U.S.: the 5-Star Rating System developed by NHTSA's New Car Assessment Program, along with the Insurance Institute for Highway Safety's (IIHS) Vehicle Rating System.

Integrating Crash Testing into the MSS

This module aims to build upon the established reliability and familiarity of the NHTSA and IIHS rating systems within the auto industry by aggregating the two to assess an AV's protective safety performance.

The value of this module in the MSS is three-fold. First, this module accounts for during- and post-crash safety performance, thereby filling in the gaps for other modules in the MSS that focus on crash avoidance performance. Second, highlighting the value of protective safety measures as part of a holistic approach to AV safety alongside preventative safety measures will help to educate the general public that AVs are not impervious to crashes. Third, the inclusion of already well-understood measures alongside more novel approaches can make the scope of creating a holistic, large-systems view of AV safety more approachable to the general public.

The NHTSA 5-Star Rating System and IHSS Vehicle Rating System

Both systems provide valuable information, and to some extent, overlap in the types of tests they conduct, but ultimately differ on their testing and scoring methods. Overall, this paper recommends using a combination of the two evaluation systems to obtain a comprehensive picture of a vehicle's structural ability to protect its occupants and other road-goers during an impact.

NHTSA 5-Star Rating System

The 5-Star Safety Ratings Program aims to provide consumers with information about the crash protection and rollover safety of new vehicles beyond what is required by federal law in the U.S.

NHTSA's rating system conducts four tests in total: (1) Frontal crash test, (2) Side barrier crash test, (3) Side pole crash test, and (4) Rollover resistance test. The results of these four tests are then aggregated to produce an overall vehicle score, where five stars represent the highest rating and one star represents the lowest. The more stars awarded means a higher level of safety.

Due to its limited budget, NHTSA has chosen to concentrate its scoring system on front and side-impact crashes that are responsible for the highest percentage of casualties. Despite these limitations, it is important to note that NHTSA is the only organization that rates rollover resistance, in addition to frontal and side crashworthiness.

IIHS Vehicle Ratings

IIHS tests evaluate both crashworthiness, how well a vehicle protects its occupants in a crash, as well as crash avoidance and mitigation technology that can prevent a crash or lessen its severity. Given that this module is only concerned with the physical structure of the vehicle, it will only consider how the IIHS tests crashworthiness.

Unlike NHTSA's 5-Star Rating System, IIHS does not aggregate the results of their tests to produce an overall vehicle rating and instead, scores each test separately, assigning each a "Good", "Acceptable", "Marginal", or "Poor" ratings.

Comparing the Two Rating Systems

The tables below summarizes the similarities and differences (Figure 3) between the two rating systems. As one covers tests that the other does not, it is recommended that both ratings are used in the evaluation of an AV's crashworthiness.

Figure 3
Similarities and differences between the NHTSA 5-Star Rating anf the IIHS Vehicle Rating

Test	NHTSA 5-Star Rating	IIHS Vehicle Rating
Frontal Crash Test	✓	✓
Overlap Front Test / Side Pole Test	✓	✓
Side Barrier Test	✓	✓
Rollover Resistance Test	✓	
Roof Strength		✓
Head Restraints and Seats		✓

Autonomous Vehicle Modular Safety Suite: A New Approach to AV Safety

3 Aggregation of Voluntary Safety Standards for AVs

Overview

Automotive development has been subjected to a variety of safety standards, both mandatory and voluntary, for several decades. However, AVs increase the complexity of standard-setting as traditional standard-setting validation methods would require the removal of the safety driver from AVs on public roads to establish a link between a given standard real-world safety outcomes. A large-systems approach mitigates these complexities by enabling standards to be, at the very least, partially validated against a number of interconnected measures of safety, as proposed by the other modules in the MSS, the majority of which do not require the removal of the safety driver.

The Safety Standard Landscape Today

Standards describe the best way to approach the development of a product or service such that the resulting product or service meets a minimum threshold of safety, reliability, and quality. These standards can be mandatory as stipulated by a government statute or regulation, or voluntary as stipulated by any number of Standard Setting Organizations (SSO) across the globe, such as the International Organization for Standardization (ISO).

Given the safety-critical nature of automobiles, automotive development has been subjected to a variety of safety standards, both mandatory and voluntary. The U.S. federal government mandates that all vehicles sold in the U.S. abide by the Federal Motor Vehicle Safety Standards (FMVSS). These standards primarily relate to crash safety, including crash avoidance, crashworthiness, and post-crash survivability. Other safety-critical aspects of automobiles are addressed by SSOs, such as ISO 26262, which covers the functional safety of electrical and electronic systems in vehicles.

The one constant across these existing automotive standards is that they aim to standardize an aspect of the vehicle, not the driver. By focusing exclusively on the vehicle, the enforcing body can validate the efficacy of these standards by inspecting the vehicle directly or by conducting evaluative tests in a lab. AVs complicate this approach because the vehicle and driver can no longer be separated; they become one and the same. Furthermore, the computerized functions of a robotic "driver" are based on machine learning algorithms, which are non-deterministic, meaning that the same inputs do not necessarily lead to the same outputs. This black-box nature of machine learning algorithms makes it extremely difficult to audit the robotic driver directly as you could with hardware, or deterministic software.

Thus, future safety standards explicitly intended to govern AVs will face inherent difficulties in validating their efficacy. In a small-systems approach, the only way to validate the efficacy of a set of safety standards is by using real-world safety data, namely the number of crashes, injuries, and fatalities on a per-mile basis, which is not only inherently unethical to gather in the early stages of deployment but would also require several years of testing before a statistically significant amount of data can be obtained for analysis.

A Large-Systems Integration of Existing and Future AV Standards

This module aims to aggregate and track existing and future AV standards put forth by various SSO. The MSS can add immediate value to the creation of these safety standards by tracking the utilization of these standards across AV companies, and mapping their utilization to safety metrics outlined in other modules in the MSS, specifically Modules 2, 4, and 6, which measure both real and simulated safety outcomes.

By evaluating the efficacy of these standards against safety outcomes, this module also aims to help SSOs identify the strengths and weaknesses of their approach quickly so that they can continue to improve the efficacy of their standards via immediate feedback loops. Additionally, this will help policymakers identify which of the AV standards contain recommended practices which are most useful for ensuring better safety outcomes, and thus should be considered in the formulating of policy.

"Utilization" can be tracked in two ways:

1. List of Uncertified Standard Adopters: AV companies that aim to adopt said standard but have not yet satisfied all the standard's requirements.

2. List of Certified Standard Adopters: AV companies that have adopted said standard, and have been certified by the SSO as meeting all the standard's requirements.

Once aggregated and centralized, mapping utilization of standards to performance measures from other modules will be a relatively straightforward task. Combined with the MSS' inclusion of proactive safety measures, insights regarding the effectiveness of a standard in ensuring better safety outcomes can begin to be gathered, prior to the safety driver being removed from the AV on public roads.

Example insights include:

- AV companies that are certified as having passed Standard X have a hard braking event half as often as AV companies that are not certified.

- AVs that are certified as having passed Standard Y were 20% less likely to experience a simulated sensor failure in the presence of fog as compared to AVs that are in the process of meeting Standard Y, and are 60% less likely than AVs that are not currently aiming to adopt Standard Y.

While this mapping is not intended to be fully indicative of a particular standard's value in ensuring safety performance, it should be directionally insightful to start identifying trends, strengths, and weaknesses across varying standards.

Additionally, simply by centralizing all existing and future standards and conveying their value and utilization in an easily digestible manner for the general public, the MSS can increase public awareness of these different safety standards, the value they have provided to the auto industry historically, and the value they will create in the AV space in the future. This could ultimately lead to more confidence in a Federal or State mandate that leverages these same voluntary safety standards as its foundation.

There will eventually be a comparable federal mandate for AVs as there is for today's human-driven vehicles with the FMVSS. To do so most effectively, consumers and policymakers should be evaluating the efficacy of proposed safety standards with as much data as possible. The MSS does not aim for this approach to displace the value of directly linking best practices with real-world data nor does it aim to fully mitigate the fact that machine learning algorithms are inherently impossible to validate. Rather, this "aggregate and map" approach will create a new data set that AV stakeholders, such as policymakers and consumers, can rely on when determining which standards are the most effective in ensuring that AVs are safe for deployment.

Autonomous Vehicle Modular Safety Suite: A New Approach to AV Safety

4 Proactive Safety Metrics

Overview

This module proposes a framework for the creation of a new classification of proactive safety metrics that can be used to showcase the positive safety attributes of AVs, beyond showcasing their potential to minimize the negative safety attributes of human-driven vehicles. These proactive safety metrics can be gathered safely because they do not require removing the safety driver from the AV on public roads.

New, Robust Metrics Need to be Developed to Evaluate AV Safety

As AVs emerge as a new, safety-critical innovation, AV stakeholders, including policymakers, consumers, and other road-goers, will need to be able to decide whether to permit AVs on public roads for testing and deployment. As with any innovation that impacts public safety, gathering sufficient data to determine whether the innovation meets an acceptable safety threshold is crucial in drafting effective policy and generating public trust in the technology. Historically, the transportation industry has relied on three lagging metrics as the core determination of safety: crashes, injuries, and fatalities. While using these metrics to compare the relative safety of AVs to human-driven vehicles is indicative and valuable, there are two fundamental issues with using these metrics alone to define the safety of AVs.

The first issue is a product of our historical reliance on these lagging metrics. By relying only on these traditional metrics, policy makers and society must willingly allow an undetermined future number of crashes, injuries, and fatalities to be caused by these autonomous systems in order to gather a statistically significant sample size of comparative data. This would likely result in widespread criticism from the media and safety advocates, which could deter policymakers from allowing these methods of data collection to be employed in their jurisdictions.

The second issue is that these historical metrics only capture a small fraction of unsafe actions that occur on the road on a regular basis. Other risky road behaviors, including tailgating, stopping in an intersection, weaving within a lane, among others, currently go undetected due to a lack of monitoring infrastructure on public roads. However, the sensor suites aboard AVs today would enable such granular data to be captured for the first time. Relying only on traditional lagging metrics would mean ignoring such data, which would otherwise be useful in establishing the relative safety of AVs as compared to human-driven vehicles.

Unlike other modules within the MSS discussed thus far, which are aggregations of existing AV safety measures, this module aims to create a novel framework that enables AV companies to create, measure, and showcase new measures of safety that do not require the removal of safety drivers from AVs. This framework fits well within the large-system structure of the MSS in that it not only brings to bear a new approach to thinking about AV safety that complements traditional approaches, but it also leverages the capabilities and features of other modules to increase its value. Particularly, this module, along with Module 5 (Human-to-Machine Comparative Driving Test), demonstrate how this large-systems approach enables synergies that make the whole greater than the sum of its parts. This will be discussed in greater detail in the following section.

The new metrics derived from this framework are designed to resolve the two aforementioned issues in that they can (1) be accurately gathered and assessed without removing the safety driver, and (2) create a new, holistic paradigm for how we should measure the safety performance of an AV. This framework will ideally lead to greater public confidence in this new technology, as well as increased comfort in the ongoing testing of AVs. The goal of this module is not to define what these proactive safety metrics should be. Rather, this module creates a simple, yet robust framework that will enable AV companies to share these new metrics with the general public and policymakers in an accessible manner.

Creating a Framework for Proactive Safety Metrics

This module proposes a framework that succinctly communicates what proactive safety metrics are and provides context as to why each metric is valuable. This framework consists of two axes. On the y-axis lies the fundamental responsibilities of the driver, whether that be human or robot. This will enable the audience to match each new metric with the driving task it is evaluating. The x-axis describes the safety performance of an AV as compared to that of a human driver. This allows the user to quickly and intuitively grasp the relative strengths and weaknesses of humans versus computers when it comes to executing safety-critical processes.

Y-Axis: The Four Fundamental Responsibilities of a Driver:

Perception: Effectively identifying all the objects, animate and inanimate, in the surrounding environment that are relevant to vehicle operation.

Prediction: Making predictions on how the relevant, animate objects perceived are going to behave on the road. For example, determining whether a pedestrian who is looking down at their phone is going to cross the road or not.

Planning: Given all the surrounding objects and their projected pathways, the driver plans their desired pathway to achieve their objective while abiding by the rules of the road.

Execution: Turning intent into action via vehicle control systems. For this task, humans require physical input controls, such as steering wheels and pedals while AVs do not; though both humans and autonomous systems now rely on drive-by-wire control systems to translate human or software input into vehicle action.

Other: Anything that falls outside the domain of the four fundamental responsibilities.

X-Axis: Measuring AV Performance with Human Performance as a Benchmark:

Not There Yet: These are tasks that both humans and AVs can accomplish, but human drivers are still superior in their accuracy or achievement. However, we expect safety-critical measures in this category to eventually transition to the "Superhuman" category described below, to parallel the advancement of AV technology over time.

Superhuman: These are tasks that both humans and computers can accomplish, but computers have already reached superior accuracy or achievement. Given the rate of technological advancement, it is unlikely that any measure in the "superhuman" category will transition back into the "Not There Yet" category.

Innate Advantage: These are fundamental properties of AVs that humans either do not have or cannot achieve.

An example of how the framework can be used is shown in Figure 4. The metrics included are for illustrative purposes only.

Figure 4
Example Proactive Safety Metrics Framework

	Not there Yet	Superhuman	Innate Advantage
Perception	Perceive and understand scenarios that have not been seen before (in the case of AVs, "seen" refers to whether the AV has been trained on the scenario data)	Hyper-accurate depth and speed detection	Having a 360 degree field of view always Abilities improve over time, rather than deteriorate as more data is collected over time
Prediction	Innate understanding of human intent Understanding of local driving styles		Ability to evaluate performance day-by-day
Planning	More flexible decision making, for example, being able to drive off the road accomplish a maneuver, thereby breaking road rules, but not putting anyone in direct harm	Higher consistency of obeying optimal positioning from other objects on the road Optimizing for minimal harm and damage in the event of a crash Safety mode is activated when the minimum confidence threshold is not met to create more decision-making time Is able to identify "escape routes" at all times to ensure that passengers and road-goers are safe	Consistency of reactions to consistent situations Perfect memory
Execution	N/A	N/A	N/A
Other	Communicate intent with external world	Superior latency from perception to execution	Free from human error (intoxicated, distracted, tired driving)

Additionally, there needs to be robust, comprehensive, and generally accepted methods of measurement and validation to back up performance claims. The framework does not mandate that AV companies share their supporting data, but it would benefit public understanding for the AV companies to share how they assess each metric. Several possible assessment mechanisms are outlined below. The list is not exhaustive, but all mechanisms should be able to be executed without removed the safety driver from the AV.

Lab-based tests: These measures can be tested in a lab, or in a controlled setting through simulated, scenario-based driving tests.

Human-to-Driver Comparative Driving Test: The Comparative Driving Test is a driving test that puts a human driver behind the wheel of an AV to drive through an area that is within the AV's Operational Design Domain (ODD), while the autonomous system runs in the background. This assures that factors such as conditions and routes are kept constant such that the safety of the choices made by the human and the autonomous system on the road can be fairly compared.

Proprietary: AV companies may develop unique safety assessment systems which they believe to be part of their competitive advantage over other players in the industry. We do not propose they share details of such proprietary methods, but AV companies should acknowledge that they conducted a proprietary assessment if they choose to report a metric within the framework.

Intrinsic: Innate properties of AVs. These cannot be measured but are a given.

Figure 5 demonstrates how we can add this additional context to the example previously shown in Figure 4, by color-coding each metric according to the assessment mechanism that was used to validate its place in the framework.

Figure 5
Color-Coded Proactive Safety Metrics Framework

Legend

Proprietary	Lab-Tested
Drivers Test Validated	Intrinsic

	Not there Yet	Superhuman	Innate Advantage
Perception	Perceive and understand scenarios that have not been seen before (in the case of AVs, "seen" refers to whether the AV has been trained on the scenario data)	Hyper-accurate depth and speed detection	Having a 360 degree field of view always Abilities improve over time, rather than deteriorate as more data is collected over time
Prediction	Innate understanding of human intent Understanding of local driving styles		Ability to evaluate performance day-by-day
Planning	More flexible decision making, for example, being able to drive off the road accomplish a maneuver, thereby breaking road rules, but not putting anyone in direct harm	Higher consistency of obeying optimal positioning from other objects on the road Optimizing for minimal harm and damage in the event of a crash Safety mode is activated when the minimum confidence threshold is not met to create more decision-making time Is able to identify "escape routes" at all times to ensure that passengers and road-goers are safe	Consistency of reactions to consistent situations Perfect memory
Execution	N/A	N/A	N/A
Other	Communicate intent with external world	Superior latency from perception to execution	Free from human error (intoxicated, distracted, tired driving)

Lastly, all of these metrics should be technology-neutral and performance-based, rather than prescriptive. In other words, a target should be set in terms of safety performance, but evaluators should be agnostic as to how exactly that target is achieved. This ensures that no particular AV company is favored arbitrarily due to their particular technological approach, in order to encourage innovation rather than stifle it.

The framework outlined creates a more holistic approach to understanding the safety benefits and limitations of autonomous technology. The goal is to reframe the conversation around safety from one that focuses on how AVs can decrease vehicle-related injuries and fatalities, to one that also highlights the positive elements AVs bring to road safety moving forward. Ultimately, the framework will educate AV stakeholders such as consumers and policymakers, enabling them to make more informed decisions on whether to permit AV testing and deployment on shared roads, as well as whether to use an AV service when it becomes available to the broader public.

Autonomous Vehicle Modular Safety Suite: A New Approach to AV Safety

5 Human-to-Machine Comparative Driving Test

Overview

With the advances in sensor technology present on AVs, we, for the first time, have the capability to gather data on a much wider range of driving behaviors that put passengers and road-goers at risk, such as tailgating, speeding, among others. This module proposes the creation of a Human-to-Machine Comparative Driving Test, a driving test that puts a human driver behind the wheel of an AV to drive through an area that is within the AV's Operational Design Domain (ODD), while the autonomous system runs in the background. This assures that factors, such as conditions and routes, are kept constant so that the choices made by the human and the autonomous system on the road can be fairly compared. This test also makes this comparison more comprehensive by taking into account previously undetectable behaviors, beyond traditional safety metrics that measure crashes, injuries and fatalities.

A Lack of Public Trust Can Prevent AVs From Achieving their Safety Potential

AVs aim to make the roads safer for all. However, as discussed in the previous section, the traditional measurements of road safety (crashes, injuries, and fatalities) are insufficient in capturing the full spectrum of vehicle safety. This is because they fail to capture other risky driving behaviors that occur without resulting in a crash, injury, or fatality, such as tailgating, weaving within a lane, stopping in an intersection, among others.

Additionally, the safety benefits of AV technology cannot be realized without acceptance and adoption of the technology by the public. It is expected that people will be uncomfortable with technology they are unfamiliar with. AV companies recognize this, investing significant resources into educating the public on AVs in an attempt to shift consumer sentiment in favor of this nascent technology. However, after years of demo rides, press releases, and media coverage, the needle has failed to move sufficiently. Public trust in AVs remain markedly low; out of 1,200 surveyed in a Bloomberg poll, 48 percent said that they would "never get in a taxi or ride-share vehicle that was being driven autonomously".

Compounding this skepticism in AV technology is the fact that human drivers tend to overestimate their driving ability. Despite the fact that more than 90 percent of crashes are the result of human error, 73 percent of U.S. drivers nevertheless consider themselves to be "better-than-average" drivers according to a 2018 AAA study. Even if AVs are statistically proven to have better safety performance than the average human driver, an overconfidence in human driving abilities could stifle adoption, as many might still be convinced that their driving is safer than that of AVs.

Comparing the Driving Performance of Human Drivers to AVs

In response to this, this module proposes the implementation of a new driving test to increase public trust in AVs and to instill in the public a more accurate, comprehensive understanding of how human drivers actually perform on the road.

Most AV test vehicles today still come equipped with traditional human driving controls, such as a steering wheel, pedals, and mirrors. The Human-to-Machine Comparative Driving Test will put members of the general public behind the wheel of an AV and allow them to drive through an area that is within the AV's Operational Design Domain (ODD), while the autonomous system runs in the background. In other words, the autonomous system will do everything it normally does, except execution. This gives the human complete control over the vehicle's physical operation but allows the vehicle to collect data on what choices the autonomous system would have made at every moment along the route.

This assures that driving performance factors, like conditions and routes, are kept constant so that a fair comparison can be drawn between the human driver and the autonomous system, while capturing the entire spectrum of driving behaviors to make the comparison all the more comprehensive.

While this entire proposal hinges upon public's willingness to participate participation in this driving test, there are a number of mechanisms that could generate meaningful participation:

1. AV companies could mandate that potential customers in early-launch cities must first participate in this driving test before utilizing their AV service, as a part of consumer education.

2. AV companies or local municipalities could incentivize participation in order to gather this valuable data.

3. Consumers or members of the public intrigued by the opportunity, or skeptical of the AV's capabilities, could volunteer to participate in the test out of individual curiosity.

Making Sense of the Comparative Data

Upon completion of the test, the AV company will be able to examine the data acquired by the onboard sensors and computer to understand how the AV and human driver performed respectively. Examples of data analyses can include but are not limited to the proactive safety metrics outlined in Module 4 such as comparing measurements of center lane drift, rate of controlled approach to a stop sign, reaction time to the appearance of an obstruction, among other metrics. While Module 4 could exist on its own to gather proactive safety metric data on AVs, its value increases when utilized in conjunction with this module because the same sensor suite can now gather the same data on human drivers for direct comparison. The helps to create additional human performance benchmarks for AV companies to leverage when claiming their system is "safer than a human".

Making Sense of the Comparative Data

The data gathered from Human-to-Machine Comparative Driving Tests can be utilized by multiple parties in the following ways:

Policy Makers and Regulators:
Policy makers require data in order to make informed and effective policy decisions. Through this test, policymakers can get accurate safety data on the average human driver, as well as that of AVs. This can help to make more accurate assessments on the relative safety of AVs by using human driving performance as a familiar and accessible benchmark.

AV Companies:
The data and analysis collected from these tests can be used to inform an AV company about its AV's performance, exposing edge cases and emphasizing strengths and weaknesses relative to that of human drivers.

General Public:
This allows the general public to gain access to unbiased data on their own safe-driving performance, and more importantly, how their performance compares to that of other human drivers and AVs. This awareness could allow them to level-set on their own driving capabilities and become more informed on the limitations human drivers face on the road. The nature of labeling it a "driving test" will encourage peak human performance, such that relative benchmarks are set as high as possible.

Given the direct comparison of their peak performance relative to that of the AV, human test-drivers gain a better understanding of all the things AVs do on the road to keep their occupants and those around them safe. This could serve to increase the public's confidence in AV technology and encourage adoption when AVs become widely available.

Overall, the Human-to-Machine Comparative Driving Test creates a first-of-its kind dataset covering the safe-driving performance of humans and AVs on public roads under controlled conditions. This data can be used by all parties in the ecosystem to better understand what it means to "drive safely" and how to accurately compare AV performance to human performance for evaluative purposes. Additionally, through a deeper understanding of how AVs aim to protect their occupants on the road, humans participating in this driving test may gain greater confidence in these autonomous systems, likely increasing future acceptance and adoption rates. One additional consideration is that this Module is highly complementary to Module 4 (Proactive Safety Metrics) in that it can help to measure and evaluate the proactive metrics put forward in Module 4. While these two modules can be used independently, together they demonstrate how the whole can be greater than the sum of its parts in terms of the value the MSS can bring to all stakeholders in the AV ecosystem.

Autonomous Vehicle Modular Safety Suite: A New Approach to AV Safety

6 Creation of Anonymized Databases of Best Practices, Safety Metrics, and Driving Scenarios

Overview

The sharing of information, along with the dissolution of data silos, is crucial for the convergence of industry practices that would allow for the development of data-driven safety standards governing Autonomous Vehicles (AVs). However, this desire for information must be balanced by the need for AV companies to protect their intellectual property (IP) and trade secrets in order to continue to foster innovation within the space. This module proposes that the MSS include anonymized databases of (1) best practices, (2) safety metrics, and (3) driving scenarios that can be accessible to both regulatory bodies and the general public. The voluntary and anonymized nature of the databases ensures that AV companies are able to protect themselves, while creating an incentive structure that promotes information sharing to enable all parties to make data-driven decisions, ultimately leading to the safer deployment of AVs.

Finding a Balance Between Data Transparency and Protecting Trade Secrets

Without concrete data on the performance and safety measures of AVs, it will be difficult for policymakers to implement effective policies and gain public trust for wide-scale adoption. However, AV companies have been wary of making their data open to the public due to IP concerns, and justifiably so as many companies view the data they have gathered as part of their competitive advantage. These conflicting positions need to be resolved in order to achieve effective and safe rollout of the technology.

Before considering where the middle-ground might lie, it is necessary to understand the merits of both perspectives. From the data-protectionist perspective, the data gathered by AV companies and the way in which they go about gathering that data can indeed be part of the competitive advantage for these companies. As self-driving algorithms rely on machine learning, the output from these algorithms is only as good as the input. Thus, being able to gather valuable data efficiently can be a key determinant in the success of any given AV company. The sharing of specific data, or the specific methods used to gather data across the industry could decrease the value of such data for these companies, and by extension, decrease the incentive for these companies to invest in valuable data collection, such as the collecting of rare edge-case driving scenarios.

From an open data perspective, the free sharing of data across the industry would serve two key purposes. First, and most importantly, it would increase the safety of AVs across the industry in the following ways:

1. AV developers could train and validate their technology on a wider, more comprehensive dataset, allowing them to avoid repeating mistakes that others have made.

2. Edge cases can be aggregated across the industry, decreasing the likelihood of any given edge case causing a disruption or incident for all AV operators

3. Sharing data can increase consistency of AV performance across different AV companies. This consistency will make it easier for AVs to understand what AVs from other companies are likely to do on the road and decrease technological confusion for consumers across companies.

Secondly, the free sharing of data across the industry would make it faster and easier to implement effective regulatory policies and safety standards in the following ways:

1. Building a shared pool of data supports the formulation of data-driven safety standards for AVs. It is difficult to come up with a safety threshold to be met if there is not a shared platform upon which to establish a scale of measurement.

2. It mitigates the risk of one company's best practices being implemented as the de facto standard across the entire industry, which would require all other players to fall in line with that company's procedures.

3. Without data sharing, it will take longer for the approaches of many of these AV companies to converge, making it inherently more difficult to formulate a set of standards that is applicable and fair to every company in the industry

The Creation and Integration of Anonymized Databases into the MSS

This module proposes the integration of three, distinct anonymized database types into the MSS to facilitate the sharing of (1) best practices, (2) safety metrics, and (3) driving scenarios to promote the open exchange of safety information for the purposes of understanding, development and assessment. Figure 6 highlights the characteristics and purpose of each database, as well as examples of the data that would be stored in each of them.

The voluntary and public structure of this module should allow the ecosystem to converge towards a level of data sharing that is comfortable for all players in the AV ecosystem. Additionally, the aggregation and integration of these distinct database types into the MSS will enhance the value of each database type beyond the value that is to be gained if these data types were to be used independently from each other. This is because while distinct data types are valuable in their own right, bringing different data types together can enable additional insights and patterns to be drawn regarding the value and efficacy of the individual data types in evaluating AV safety. In this way, the large-systems approach enables intra-module synergies between the data types discussed within this module, as well as inter-module synergies across the MSS.

The anonymized nature of the databases should alleviate a number of the concerns AV companies may have around sharing data such as having their performance judged prematurely by other stakeholders based on the data they publish. This will ideally increase the willingness of AV companies to share data in the first place, while still providing a high level of value to the rest of the ecosystem.

Figure 6

Best Practices, Safety Metrics, and Driving Scenarios

Type	Not there Yet	Superhuman	Innate Advantage
Best Practices	High-level guidelines or procedures to follow when encountering common elements in specific driving scenarios *(How to approach the problem)*	Allows AV companies to converge on how to best tackle common challenges in autonomy. Once converged, safety standards can be formulated and implemented with greater ease. Allows the driving environment to be more predictable for AVs and human drivers, thus decreasing complexity and confusion on the road.	In response to X number of pedestrians crossing the vehicle's pathway (element), the vehicle slows down to allow for increased decision-making time (practice)
Safety Metrics	Frequentist measures describing how often an AV successfully handles a common element across a variety of driving scenarios OR Performative measures describing common driving characteristics *(How to measure performance)*	Create a holistic understanding of how different AVs are performing when faced with different types of driving scenarios, which can help determine the threshold an AV needs to cross before it is deemed safe enough to be deployed on public roads.	Across all different driving scenarios, the AV avoids hitting pedestrians in crosswalks 999,998 times out of 1,000,000. The AV achieves a standard deviation of 4.2inches from the center of the lane over 1,000,000 miles traveled.

Figure 6

Best Practices, Safety Metrics, and Driving Scenarios (continued)

Type	Not there Yet	Superhuman	Innate Advantage
Driving Scenarios	Raw sensor data structured in the form of problems the AV is to solve or has solved; one specific visual instance or realization of conditions on the road (How to showcase the situations encountered)	Fills in unique data gaps by providing AV companies access to a pool of unfamiliar edge cases and driving conditions Build more robust systems as AV companies can test and validate their technology on a more comprehensive dataset.	AV approaching an intersection while a pedestrian approaches the crosswalk from the sidewalk while staring down at their phone A motorcycle splitting lanes is approaching from the right-rear in rush hour at dusk

It is important to note that AV companies can voluntarily opt into all, some, or none of these databases. Furthermore, there is no requirement for an AV company to share all the data it has gathered to a specific database. As an example, an AV company can opt into the Best Practices database but choose to only share 80 of their 100 best practices. Additionally, each database should be managed by an independent, third party to ensure full oversight of the data, protection from cybersecurity attacks, and shielding of misappropriation by any one AV company or government body.

Using the Databases Individually (Small-System):

The value that can be extracted by each of the constituents from each of the databases are as follows:

Policymakers

Best Practices: Even without direct attribution, policymakers can get an idea of which best practices might in fact be representative of industry standards based on the number of companies that share same or similar ones. For example, if 80 percent of the participating companies share a best practice, it is worthy to investigate whether such a practice holds legitimate merit in bettering an AV's safety performance.

Safety Metrics: Policymakers can get a sense of the type of tasks AVs are proficient in, as compared to tasks that still require significant development.

Driving Scenarios: Policymakers can use driving scenarios to build a localized simulation driving test to evaluate AVs that are applying for testing permits in their jurisdictions.

AV Companies

Best Practices: By learning what their peers are doing, AV companies can decide whether a best practice they have not yet considered is worth implementing.

Safety Metrics: By tracking the performance of their peers, AV companies can get a sense of where certain aspects of their tech stacks are in terms of maturity and capability.

Driving Scenarios: AV companies can use shared data to fill in unique data gaps to avoid making the same mistakes as their peers, or to expand their edge case catalogue more efficiently. However, in order to create a market for this data there needs to be an incentive structure where AV companies can only get access to a greater proportion of the data pool if they share a larger volume of data, or share more unique data points.

General Public

Best Practices: This will help the general public better understand how AVs are developed, potentially generating more confidence and trust in these systems.

Safety Metrics: Members of the public can get a sense of the type of tasks AVs are relatively proficient in, as compared to tasks that still require significant development.

Driving Scenarios: No value on its own.

Combining the Databases to Leverage on Synergies (Large-System):

While each database has value on its own, significant value can be created if datasets within different databases are linked to each other or to other modules in the MSS.

For one, by linking a company's Best Practices dataset with their Safety Metrics dataset, policymakers and regulators can get a more complete picture of which best practices lead to the better safety outcomes, and such findings could inform policymaking. By linking a company's Driving Scenario database with their Safety Metrics database, policymakers and regulators could get a sense of quantity and quality of simulated and real-world scenarios required to reach a certain level of safety for each metric, which could inform policymaking and the setting of safety standards as well.

It is important to note that even when linked, the datasets can remain anonymous; users would only need to know that a certain set of best practices or driving scenarios leads to a certain set of safety outcomes, without having to know the data source.

However, while this module focused exclusively on anonymized databases because they have the highest likelihood of attaining industry buy-in, any AV operator could always decide to deanonymize their data. This could be for a number of strategic reasons; for example, an AV company might wish to showcase what they believe to be superior capabilities by releasing their safety metric data, or demonstrate that there is able to handle an impressive level of complex driving scenarios, in order to gain trust from other stakeholders, including potential consumers and relevant policymakers.

Autonomous Vehicle Modular Safety Suite: A New Approach to AV Safety

Concluding Remarks

Autonomous technology is expected to drastically improve the safety, sustainability, and mobility of our transportation system, introducing a paradigm shift to the way we move and transport goods. However, as there were risks in the transition from horses to automobiles a century earlier, there are risks involved in the current transition from human-driven vehicles to autonomous ones.

How can we as a society attain the promised benefits of this new technology as quickly as possible, while mitigating the life-threatening risks that present themselves along the way? The short answer is that no one yet knows, and it must be acknowledged that there is no certainty without the benefit of hindsight; the aviation industry and its history of reactive regulation creates precedence for accepting this reality. While this uncertainty is unsatisfying, technological advancement cannot stand still because of it.

Acknowledging that creating a cohesive and inclusive approach to safety is the key to accelerating AV development, the MSS offers a new way of thinking about AV safety. It aims to create a large-system framework that takes into account traditional and innovative, quantitative and qualitative safety measures through the aggregation and integration of smaller system solutions. With the MSS, the whole is greater than the sum of its parts.

To achieve this vision, collaboration is crucial. This white paper urges automobile manufacturers, technology developers, and relevant government agencies to come together to build systems, share knowledge, and establish protocols. If the existing silos are to remain, the development of a cohesive safety framework like the MSS cannot move forward, and the advancement of AV technology will stall.

This proposed approach also highlights the need for agility, given the technology's nascency. The MSS is built to become stronger with new insights. As new modules are added, the convergence towards a set of best safety practices is accelerated through the validation or invalidation of existing approaches. It recognizes that AV safety needs to adapt as autonomous technology evolves over time.

Important questions still need to be addressed before the MSS can begin to effect positive change on the AV industry. Nevertheless, this white paper represents the first step towards developing a comprehensive, agile, and inclusive large-systems framework for the improved understanding and assessment of AV safety for all.

Autonomous Vehicle Modular Safety Suite: A New Approach to AV Safety

Bibliography

AAA Automotive. 'Many Americans Remain Afraid of Fully Self Driving Vehicles', 8 May 2019. https://www.aaa.com/autorepair/articles/many-americans-remain-afraid-of-fully-self-driving-vehicles.

Anderson, James M., Nidhi Kalra, Karlyn D. Stanley, Paul Sorensen, Constantine Samaras, and Tobi A. Oluwatola. 'Autonomous Vehicle Technology: A Guide for Policymakers', 22 March 2016. https://www.rand.org/pubs/research_reports/RR443-2.html.

Blumenthal, Marjory, Laura Fraade-Blanar, Ryan Best, and J. Luke Irwin. Safe Enough: Approaches to Assessing Acceptable Safety for Automated Vehicles. RAND Corporation, 2020. https://doi.org/10.7249/RRA569-1.

Boudway, Ira. 'Americans Still Don't Trust Self-Driving Cars, Poll Shows'. BloombergQuint, 19 May 2020. https://www.bloombergquint.com/business/americans-still-don-t-trust-self-driving-cars-poll-shows.

Edmonds, Ellen. 'More Americans Willing to Ride in Fully Self-Driving Cars'. AAA Newsroom, 24 January 2018. https://newsroom.aaa.com/2018/01/americans-willing-ride-fully-self-driving-cars/.

Kratsios, Michael, and Elaine Chao. 'Ensuring American Leadership in Automated Vehicle Technologies: Automated Vehicles 4.0'. United States Department of Transportation, January 2020.

Strickland, Grace, and John McNelis. 'Autonomous Vehicle Reporting Data Is Driving AV Innovation Right off the Road'. TechCrunch (blog), 4 August 2020. https://social.techcrunch.com/2020/08/04/autonomous-vehicle-reporting-data-is-driving-av-innovation-right-off-the-road/.

Wiggers, Kyle. 'Aurora Urges Autonomous Vehicle Industry to Adopt Better Safety Metrics'. VentureBeat (blog), 24 January 2020. https://venturebeat.com/2020/01/24/aurora-urges-autonomous-vehicle-industry-to-adopt-better-safety-metrics/.

AV Combinator

April 2021

Designed by Jenn Hu

www.avcombinator.com