SUBMITTED ELECTRONICALLY VIA REGULATIONS.GOV

Dr. Steven Cliff
Acting Administrator
National Highway Traffic Safety Administration
1200 New Jersey Avenue S.E., West Building
Washington D.C. 20590-0001

April 1, 2021

Re: NHTSA Advance Notice of Proposed Rulemaking, Framework for Automated Driving System Safety, NHTSA Docket No. 2020-0106, 85 Fed. Reg. 78058 (December 3, 2020)

Dear Acting Administrator Cliff:

Intel Corporation ("Intel") appreciates this opportunity to provide comments to the National Highway Traffic Safety Administration's ("NHTSA" or "Agency") advance notice of proposed rulemaking ("ANPRM") for the development of a framework for Automated Driving System ("ADS") safety. This framework would objectively define, assess, and manage the safety of ADS performance while ensuring the needed flexibility to enable further innovation.

In our comment submission, Intel proposes a two-step process that we believe best positions the United States ("U.S.") to be a global leader in Automated Driving Systems ("ADS"): near-term definition by NHTSA of performance-oriented metrics for ADS to complement development of Federal motor vehicle safety standards ("FMVSS") for ADS in the longer-term. Intel, through our subsidiary Mobileye, has extensive experience in advanced driver assistance systems ("ADAS") and believes our proposed approach is well-aligned with NHTSA's desire to provide ADS safety without hampering innovation. Our proposal provides clear regulatory requirements in the near-term while focusing the Agency's limited resources on regulating perception and planning. We strongly recommend NHTSA specify a vehicle level performance requirement that ADS-equipped vehicles shall perform at least as good as human drivers, defined as the Mean Time Between Failure

(“MTBF”), and require a technology-neutral, transparent approach to planning based upon a formal safety model. Such an approach means that the only contributing element to the vehicle level failure rate (e.g., the MTBF) would be the perception system. Additionally, NHTSA should prescribe values for reasonably foreseeable assumptions of other road user behavior for use in safety models like Responsibility-Sensitive Safety (“RSS”) used within ADS-equipped vehicles. These values could readily be defined based upon existing NHTSA research, and a framework for a subset of these parameters is nearing finalization in a widely attended industry-led standards effort, Institute of Electrical and Electronic Engineers (“IEEE”) 2846[1]. We provide a high-level background on our approach below, followed by answers to each of the questions posed in this proceeding, as well as an appendix with additional detail about formal safety models.

Intel commends NHTSA for its leadership, support, and public education relating to the development of vehicles equipped with ADS. NHTSA has published numerous research reports, guidance documents and rulemakings, and taken actions to remove regulatory barriers to innovative safety technology and facilitate the safe testing of self-driving technology. Additionally, NHTSA launched activities, such as the Automated Vehicles Transparency and Engagement for Safe Testing (“AV TEST”) Initiative, to promote public engagement and knowledge sharing about the safety of automated driving systems. These efforts, among many others, have established a foundation upon which to create regulatory pathways to deploy ADS-equipped vehicles at scale in the United States (“U.S.”).

Intel Technology Leadership

Intel, through its subsidiary Mobileye, is a key stakeholder in the development of ADS technology. We are a market leader in automation systems for driver assistance, with over 60 million vehicles on the road around the world today. This foundation of leadership in ADAS makes us uniquely knowledgeable about the life-saving

[1] https://sagroups.ieee.org/2846/.

capability of vehicle automation technologies. Through our Mobileye division, Intel expects to be a global leader in the delivery of Self-Driving Systems as both a supplier, and an operator of direct to consumer Mobility Services around the world. We are positioned to make automated driving a reality, and we have the collective depth and breadth of experience, talent, technology, and resources to deliver safe and scalable automated vehicle ("AV") solutions.

Intel welcomes the opportunity to discuss our proposed framework for ADS safety with NHTSA. Our approach is technology-neutral, data-driven, and creates a pathway for timely deployment of ADS-equipped vehicles at scale. We also believe that our framework is consistent with the Motor Vehicle Safety Act and coincides with the Agency's goals to objectively define ADS safety while also ensuring flexibility for new safety innovations.

Sincerely,

A. Shashua

Prof. Amnon Shashua
President & CEO,
Mobileye, an Intel company
Senior Vice President, Intel Corporation

Question 1. *Describe your conception of a Federal safety framework for ADS that encompasses the process and engineering measures described in this notice and explain your rationale for its design.*

Intel's Proposed Framework for ADS Safety

Intel proposes a two-step approach. In the near-term, it is critically necessary for NHTSA to provide clear guidance to industry on the required performance-oriented metrics for ADS performance prior to deployment in order to enable commercial deployment at scale in a timely manner. For the long-term, the framework could address the traditional FMVSS approach.

Our proposal for a framework for automated vehicle safety incorporates the following elements for NHTSA's consideration:

(1) Near-term performance-oriented metrics for ADS safety;

(2) Defining a vehicle level failure rate for ADS performance compared to human drivers;

(3) Defining values for reasonably foreseeable assumptions about behavior of other road users; and,

(4) Technology-neutral process and engineering measures, including formal models to define what constitutes driving safely for an ADS.

<u>Near-Term Performance-Oriented Metrics for ADS Safety</u>

We propose that NHTSA should expeditiously establish performance-oriented metrics for ADS safety. While we strongly support the traditional FMVSS approach, we understand that regulatory standards will take time and are likely a longer-term effort. However, clear performance-oriented metrics for ADS are needed expeditiously. It is critically necessary for the Agency to expeditiously provide clear guidance to industry on performance-oriented metrics for ADS-equipped vehicles prior to deployment in order to enable commercial deployment at scale of this important safety technology. These actions will facilitate the safe deployment of

ADS-equipped vehicles in a timely manner, in accordance with societal expectations on the safety of ADSs, while also preserving the nation's technological leadership.

Intel appreciates NHTSA's concerns about whether ADS technology may be premature for regulation and that meaningful data is still needed to analyze and determine how and which aspects of performance require regulation. We believe that performance data, along with Process and Engineering measures, provide meaningful information to prove the efficacy of ADS-equipped vehicles and such data is available today. Other countries, including France[2] and Germany[3], have recognized this and are introducing regulation in this area to support commercial deployment of ADS-equipped vehicles by as soon as 2022. Due to the tremendous innovation in ADS technology, we estimate that AVs will be ready for commercial deployment within this timeframe.

Establishing performance-oriented metrics from NHTSA will not hinder innovation, rather it will help avoid stymying the industry. Performance metrics by definition are technology-neutral as they specify a required level of performance, not an implementation. Distinct performance-oriented metrics will define the required level of safety for all ADS-equipped vehicles, independent of their technical differences, allowing for an apples-to-apples comparison across the entire industry for ADS testing. While the Agency collects and evaluates data obtained through real-world use based on near-term performance-oriented metrics for the development of longer-term FMVSS for ADS performance, developers and manufacturers could still pursue safety innovations and novel designs in ADS technologies.

Summary of Proposed Framework

Intel agrees with NHTSA that sensing, perception, planning, and control are the core elements of an ADS. However, we suggest that of these core elements, perception and planning should be the focus of the Agency's safety assurance framework. The foundation of our framework is a combination of process measures such as

[2] https://www.ecologie.gouv.fr/sites/default/files/20171_strategie-nationale-vehicule%20automatise_eng_web.pdf.
[3] https://www.bmvi.de/SharedDocs/EN/Articles/DG/automated-and-connected-driving.html.

ISO 26262 and ISO 21448 along with engineering measures, such as formal models for decision making, which will provide NHTSA and the public with confidence and accountability that an ADS is safe-by-design.

As illustrated in Figure 1, a specific performance target for the vehicle level failure rate, the Mean Time Between Failure (MTBF), is defined. Failure is defined to be an at-fault crash. Human drivers today crash on average every 500,000 miles.[4] Given an average speed of 10 mph that would translate to a MTBF (e.g., a crash) once every 50,000 hours. We believe that NHTSA should require ADS-equipped vehicles to have an MTBF of at least that of human drivers.

While it is true that there may be integrity failures in the hardware ("HW") or software ("SW") of an ADS; such faults are sufficiently mitigated by conformance with ISO 26262. The other source of potential failure is an algorithmic failure. In an ADS, there are two sources of potential algorithmic failure: the perception or planning systems.

[4] https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813060.

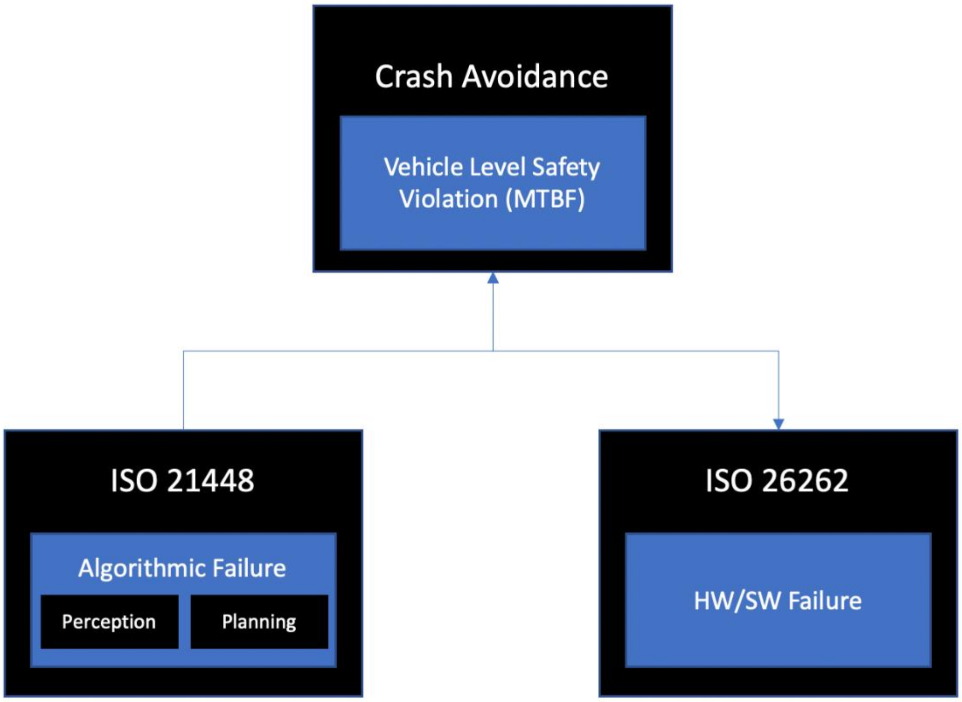

Figure 1: Decomposition of Vehicle Level Performance Target

Conformance with ISO 21448 can help identify sources of algorithmic failures. Further, if the contribution of planning can be shown to approach infinity, then the vehicle level failure rate will roughly be equivalent to the perception system failure rate.

To understand why, consider that the number, kind, size, or type of sensor is irrelevant to what really matters: the accuracy of perception. Perception is a critical function of ADS safety because in order for an ADS to operate, it must accurately perceive the surrounding environment. This includes other vehicles, pedestrians, or infrastructure around the AV, and for each detected object, its position and speed. A critical failure in the perception system can lead to a crash (e.g., if the ADS fails to perceive a stopped car in front of the AV, it will crash into the front car). Control theory is well understood and covered by existing regulations so should not be a material contributor to the overall vehicle failure rate.

Planning, or often called the “driving policy” by industry, is where the ADS must decide what to do next once it has an accurate understanding of its surroundings. While perception deals with understanding the present state of the world, planning requires “what happens next” predictions. Furthermore, the prediction of the future is not constant but depends greatly on the decisions the ADS will make.

A failure in either perception or planning could lead to a vehicle level failure (i.e., crash). Yet there are critically important differences between perception and planning. Perception can be subject to defined regulatory requirements for accuracy – there is a clear “ground truth” to what the perception system should have detected. For example, if there is a car in front of the AV and the perception system fails to recognize it, clearly the perception system has failed. As a verification check, two humans will provide the same answer when asked if there is a car in front of the AV or not. Therefore, one can easily assess the performance of the perception system by calculating how often it makes an error. A better perception system is clearly linked to a safer ADS as the fewer perception mistakes the system makes, the “safer” the ADS has the potential to be.

However, in planning, driving decisions are based on “what happens next” predictions, which do not have a “ground truth”. Two different human drivers could provide two different answers as to what the right driving decision should be in different scenarios such as when considering whether to give right of way or not to another vehicle. Thus, in planning, there is not a clear definition of what constitutes a failure, so decisions are open to interpretation and judgement of the situation. As the saying goes, hindsight is 20/20, which is apropos when considering whether a driver made the “right” decision, as such judgement is often made after the fact when a crash has happened. Unlike perception, making a planning system “safer” may not always lead to a “better” ADS. Consider that the safest vehicle is the one that never leaves the garage. An overly conservative ADS might be considered to be safer by some, but it may contribute negatively to traffic congestion, impact the behavior of other drivers (going around a slower moving vehicle), or fail to deliver a passenger to their destination. Planning, even for human drivers, foregoes absolute safety (where humans never drive) and instead involves making decisions based upon reasonably foreseeable risk, which is what human drivers do every time they get behind the wheel.

The foundational philosophy behind understanding what it means to “drive safely” is defining clear boundaries on the reasonable worst-case behavior of other road users. By incorporating reasonably foreseeable assumptions about the worst-case behaviors of other road users, an ADS can be programmed to minimize risk while balancing safety and usefulness in accordance with societal expectations.

Formal models, like Responsibility-Sensitive Safety[5] (RSS), which fully embody this philosophy have been formally verified to be safe-by-design, and as a result can ensure that the ADS will never have a lapse in judgement and will never make an incorrect driving decision within the bounds of the reasonably foreseeable assumptions about other road users embodied in the model.

Critically, this means that when utilizing a formal model like RSS to implement the planning function of the ADS, the contribution from planning to the vehicle level failure rate (the MTBF) approaches infinity. As a result, assuming there are no other mechanical or infrastructure failures, only perception failures would be the primary contributor to the overall vehicle MTBF, which significantly simplifies the safety assurance challenge, and the scope of regulation needed. Furthermore, perception systems can be efficiently tested offline to ensure performance to the defined MTBF target.

In planning systems, foreseeable risk is bounded by making reasonably foreseeable assumptions about the behavior of other road users. Regulations need to formally define the values for reasonably foreseeable risk for the ADS planning function. This is similar to a regulator selecting a speed limit for a new road. The value (e.g., speed limit) represents the level of reasonably foreseeable risk for that roadway. Reducing the maximum speed limit to 10 miles per hour on all roads in the United States would have an incredibly positive impact on safety. However, doing so would not be practicable and the negative impacts on the efficiency of the

[5] https://www.mobileye.com/responsibility-sensitive-safety/.

transportation system would be severe. Thus, a balance is achieved by setting a speed limit that accounts for reasonably foreseeable risk for a given roadway.

Note also that in order to ensure safety-by-design, it is critically necessary that the capabilities of the ADS be sufficient to handle immediate response reactions to reasonably foreseeable situations within a given operational design domain ("ODD"). For example, in a multi-lane interstate highway, a common emergency scenario is the cut-away revealing a disabled object or vehicle ahead. If a SAE Level 3-5[6] ADS was designed without side view perception, it would be unable to change lanes to avoid this reasonably foreseeable situation (i.e., disabled object/vehicle) and would have no choice but to alert the human occupant who would be unable to avoid a crash within the fractions of a second from the takeover alert. NHTSA should take special care to ensure that commercially deployed ADS-equipped vehicles – in particular SAE Level 3 ADS-equipped vehicles – sufficiently consider all reasonably foreseeable behaviors from other road users in the defined ODD and is able to handle, under automation, all necessary immediate reaction emergency situations. The real-world impact of this critically important safety-by-design goal is that the capabilities of a SAE Level 3 ADS cannot be materially different than a SAE Level 4 ADS operating on the same roadways.

In the following sections we provide further information about the elements of our proposed framework.

Defining a Vehicle Level Failure Rate for ADS Performance

NHTSA should establish vehicle level performance targets for the ADS that define the desired level of safety compared to human drivers. Utilizing the framework proposed in our response, Intel believes that it is very achievable to demonstrate ADS performance to a specified failure rate, which we refer to as the Mean Time Between Failure (MTBF), where a failure is defined as an at-fault crash by the ADS. At the most basic level, one can consider the probability of a crash in an hour of driving. It is reasonable then to set the MTBF for an

[6] https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic.

ADS to be at least the level of human drivers.

NHTSA data sets such as the National Automotive Sampling System (NASS) General Estimates System (GES) and pre-crash scenario typology studies provide a data driven path to understanding current human driver performance in different situations so that ADS-equipped vehicles can be expected perform with at least the same level of safety in the same situations; in other words that the MTBF of the ADS-equipped vehicle is at least as good as the MTBF for human drivers in the same scenario. We believe that NHTSA should require the vehicle level failure rate (or MTBF) for an ADS-equipped vehicle to be at least as good as a human driver, and ADS developers should provide evidence supporting their achievement of this performance target. By setting this performance target for an ADS-equipped vehicle, NHTSA can provide critical clarity on the expected performance of an ADS-equipped vehicle compared to human drivers, and industry can provide evidence as to how the MTBF has been met.

Defining Values for Reasonably Foreseeable Assumptions About Behavior of Other Road Users

The ADS planning function is a critical element of any ADS safety framework. As expressed by NHTSA, the planning function is equivalent to the part of the brain of the ADS responsible for decision-making. The safety goal for an ADS for the planning function should be to always make safe driving decisions; to achieve this, assumptions must be made by the ADS about the potential behavior of other road users. Just as human drivers need to make assumptions about the expected maximum braking performance of the vehicle they are following, or the possible maximum acceleration of a pedestrian on a sidewalk next to the road, an ADS must also make these same assumptions in order to operate in the real world. Then, within the boundaries of these reasonably foreseeable assumptions, the ADS should always endeavor to make the correct decision. The Responsibility-Sensitive Safety (RSS) model fully supports this philosophy, including supporting all applicable driving scenarios including right-of-way, occlusion, unstructured road scenarios and much more. RSS additionally provides clarity on what the ADS should do if a collision is imminent but can be safely avoided.

An industry-led standardization effort, IEEE 2846, is standardizing a subset of the reasonably foreseeable assumptions utilized in RSS. IEEE 2846 provides a framework within which NHTSA can define values for reasonably foreseeable assumptions about road user behavior that can be incorporated into industry safety models used in planning like RSS. Naturalistic driving studies can provide excellent guidance on what is reasonably foreseeable about the behaviors of other road users. So long as the other road users behave within the reasonably foreseeable assumptions, formal safety models like RSS make sure that the ADS will make decisions that will contribute positively to the safety goal (e.g., do not crash). While RSS goes further than IEEE 2846 by describing clear rules for right-of-way and evasive maneuvers in imminent collision scenarios, Intel believes IEEE 2846 is an excellent place for the regulator to begin providing clarity on the expectations of ADS performance.

IEEE 2846 is a consensus-based technical standard being developed by industry leaders such as Aurora, Daimler, DENSO, Ford, Honda, Motional, Stellantis, Volkswagen, Waymo, plus UMTRI, NIST, and dozens more, which describes the minimum set of reasonable assumptions used in foreseeable scenarios that shall be considered in safety-related models that are part of an ADS's planning element. In recognition of international support for IEEE 2846, formal liaison relationships have been established with SAE-ITC's On Road Automated Driving (ORAD) Committee and ISO TC22 / SC32, which is responsible for the development of ISO 26262 and ISO 21448. Aligned with our belief that the safety of automated vehicles should not be a point of proprietary differentiation but common ground for all of the automotive industry, Intel contributed its technology-neutral, transparent formal safety model, RSS, to the IEEE 2846 working group for inclusion in this consensus-driven industry safety standard. The outcome of this effort provides NHTSA with a valuable framework to define values for parameterized assumptions about the behavior of other road users.

Technology-Neutral Process and Engineering Measures

Intel agrees with NHTSA's differentiation of Engineering from Process measures, as both play important but different roles in the design, development, and deployment of automated vehicles. Process measures ensure good systems engineering practices are employed in the development of the ADS, while engineering measures provide safety-by-design technical assurance. We recommend that NHTSA support technology-neutral process and engineering measures, specifically, ISO 26262 and ISO 21448 for process measures, and recommend a formal safety model like RSS for the ADS planning function for engineering measures.

Of the process measures described in NHTSA's ANPRM, Intel believes that those defined in ISO 26262 and ISO 21448 offer industry-leading best practices for functional safety while still allowing flexibility and innovation; NHTSA has recognized these criteria are important considerations for a regulatory framework for ADS safety. As shown in Figure 1 above, ISO 26262 defines methodologies for defining, analyzing and mitigating potential random HW faults and systematic SW issues within the ADS, while ISO 21448 provides a framework to ensure the soundness of algorithmic approaches including minimizing the risk of any potential misuse.

The combination of ISO 21448 and ISO 26262 process measures during development enables ADS developers to deliver an expected rate of safety (i.e., acceptable rate of safety goal deviation) based on exposure to specific traffic situations within an ODD to more accurately describe whether an ADS is ready for its intended deployment based on its design. Additionally, these process measures can be adapted to a broad range of technologies. We propose ISO 26262 and ISO 21448 as the best combination of process measures to ensure a complete and correct analysis of the design and possible hazards present in an ADS's implementation.

For engineering measures, we strongly recommend that NHTSA consider requiring a formal model for the planning function of an ADS, consistent with NHTSA's suggested approach to "require vehicles to be programmed to drive defensively in a risk-minimizing manner in any scenario within their ODD". Formal models like RSS have been widely adopted by industry and leading safety standards organizations. Formal model approaches to safe ADS decision making are technology-neutral and provide critical transparency of the decision-making logic of an ADS, a criterion that NHTSA has highlighted is important for creating public trust. Formal models do not rely on opaque, proprietary artificial intelligence ("AI") approaches which may not be explainable to the public; rather, formal models like RSS allow regulators and the public the ability to understand and test the performance of the ADS, and to also figure out what happened and why if a collision involving an ADS occurs. Additionally, formal models can provide certified verifications of correctness to achieve safety-by-design, and offer a conceptual guarantee that if all road users were to conform to the safety model's assumptions about the reasonably foreseeable behaviors (and there were no other infrastructure or mechanical failures) then there should be no crashes. Crucially, adoption of a formal model for planning means that the primary remaining source of possible failures (e.g., the MTBF) is limited to the perception system – significantly reducing the scope and challenge of verifiable safety assurance.

Safety-by-design approaches such as formal models, long proven in other industries like avionics[7], are being widely adopted in the ADS industry to provide certainty in decision-making within the bounds of what is reasonably foreseeable. The certainty provided in planning by models like RSS eliminates the need for an ADS to perfectly predict human behavior (which is impossible). Instead, the ADS considers what is reasonably foreseeable – a standard enshrined in leading international regulations such as the United Nations Economic Commission for Europe ("UNECE") automated lane keeping systems regulation[8] which states that the ADS "shall not cause any collisions that are reasonably foreseeable and preventable". Given values for reasonably foreseeable assumption parameters in a formal model like RSS, the planning function of the ADS can ensure it will always make safe driving decisions, achieving an MTBF of near infinity within the boundaries of the values for reasonably foreseeable assumptions. As mentioned previously, IEEE 2846 provides a framework within which NHTSA can define a subset of the values for reasonably foreseeable assumptions about the behavior of other road users that should be considered in formal models like RSS used within an ADS.

Importantly, NHTSA can change the values set for the reasonably foreseeable assumptions about the behaviors of other road users over time if needed. In addition, different assumptions could be made for wet versus dry

[7] https://link.springer.com/chapter/10.1007/978-3-642-34281-3_2.
[8] https://unece.org/sites/default/files/2021-03/R157e.pdf.

driving conditions, as well as for rural versus urban driving. The regulator can dynamically adjust the desired level of ADS performance if needed, without hampering innovation. Within the bounds of the chosen parameters, nothing in the ADS needs to change – the ADS developer can make necessary enhancements without compromising safety. Performance monitoring of the ADS can be easily done by understanding the extent to which the formal model's safety envelopes were violated by either the ADS or other road users, giving the regulator a transparent and technology-neutral way to assess and adjust the real world performance of an ADS.

In summary, our proposal for a framework for automated vehicle safety incorporates the following elements for NHTSA's consideration:

(1) Near-term performance-oriented metrics for ADS safety;

(2) Defining a vehicle level failure rate for ADS performance compared to human drivers;

(3) Defining values for reasonably foreseeable assumptions about behavior of other road users; and,

(4) Technology-neutral process and engineering measures, including formal models to define what constitutes driving safely for an ADS.

For further details on formal models like RSS, please refer to the Appendix following our responses to the questions.

Question 2. *In consideration of optimum use of NHTSA's resources, on which aspects of a manufacturer's comprehensive demonstration of the safety of its ADS should the Agency place a priority and focus its monitoring and safety oversight efforts and why?*

We agree with NHTSA that sensing, perception, planning, and control are the primary functions to consider for determining ADS safety. However, we recognize that the Agency's resources are not infinite, and it must carefully allocate limited resources to focus on the elements that are of most importance – perception and planning. The kind of sensor used in an ADS is not relevant to safety so long as the output of the perception function is accurate. Additionally, the planning function is essential to ensure that an ADS always makes safe driving decisions. Control theory is well understood by the industry and does not contribute materially to vehicle level failures, so additional rules are not needed. Therefore, NHTSA should take particular care to evaluate the perception and planning functions of the ADS to assess a manufacturer's comprehensive demonstration of its safety.

As the planning element of an ADS should never have a lapse in judgment, probabilistic-based approaches for planning should be discouraged. How can NHTSA know if an ADS just got lucky during a performance test or if it passed for the right reasons, by design? Formal model-based approaches like RSS can help the Agency avoid this outcome and provide transparency and certainty which the public expects for ADS decision-making. The effectiveness of formal safety models has been proven for decades in other industries, such as avionics. By requiring manufacturers to adhere to a risk-minimizing approach for the ADS planning function, NHTSA can be confident that the planning element will not materially contribute to the vehicle failure rate. NHTSA can efficiently evaluate the performance of the planning function in simulation by recreating known pre-crash scenarios and assessing whether the planning function (given perfect perception within the simulation environment) can in fact avoid the known crash types. ADS developers can alternatively provide sufficient evidence of their own simulation tests against the defined pre-crash scenario types for NHTSA to evaluate the performance results.

NHTSA can then focus primarily on monitoring and overseeing the accuracy of the ADS perception function. ADS perception can be measured through efficient offline tests or by tracking real world performance events to determine if the vehicle meets the level of performance defined by the regulator (e.g., the MTBF).

Question 3. *How would your conception of such a framework ensure that manufacturers assess and assure each core element of safety effectively?*

Our proposed framework optimally reduces the assessment burden for manufacturers by eliminating the need to drive tens of billions of hours in the real world to gather sufficient statistical evidence, while still providing the safety assurance of the ADS planning function. Evaluating a formally verified, technology-neutral safety model is a straightforward verification process that can easily be performed with existing industry tools and methods, as has been done in other industries such as avionics for decades. For example, verification of the planning element can be done using simulators that re-create variations of known crash types. As a simulation environment provides perfect perception for the planning element, it is an excellent way to isolate and test the correctness of the planning element.

As stated previously, when considering the vehicle level failure rate if a formal model is used for planning (and there are no other mechanical failures in the vehicle or infrastructure), the primary contributor to the overall MTBF is limited to perception. Manufacturers can easily evaluate the performance of its ADS perception element using off-line data-replay tests to assess at what rate a perception error occurs in the ADS that could lead to a crash, or provide evidence of successful operation in the real world as a measure of hours of operation without a perception failure that led to a safety incident (e.g., the MTBF).

Question 4. *How would your framework assist NHTSA in engaging with ADS development in a manner that* *helps address safety, but without unnecessarily hampering innovation?*

Our proposed framework can achieve a high level of safety across the entire industry by prescribing clear performance metrics while also providing flexibility to allow for significant innovation by manufacturers to determine how to meet the NHTSA defined performance metrics. Our proposed framework has a two-phased approach. In the near-term, NHTSA would expeditiously establish performance-based metrics, such as a minimum MTBF of the ADS-equipped vehicle to be at least that of human drivers. By providing these performance metrics, each ADS developer will have a common understanding of the requirements rather than leaving it up to each ADS developer to decide on their own what failure rate is acceptable, thereby enabling ADS safety and facilitating broader public trust in ADS technologies. A lack of common guidance until FMVSSs are developed in the longer-term could stymy ADS technology investment and the benefits they promise for our society. The lack of minimum performance guidelines also enables bad actors to deploy potentially very unsafe ADS-equipped vehicles on the road.

Our proposed approach is technology-neutral and does not specify or require any particular form, number, kind or type of sensor, or product from any one supplier. Instead, our framework focuses on the accuracy of the output of the ADS perception function. Much like it doesn’t matter what color eyes a human driver has, what is essential is how well the human driver perceives their driving environment.

A formal safety model for objectively assessing the ADS planning function, using reasonably foreseeable assumptions about the behavior of other road users defined by the regulator, is also technology-neutral and can be combined with any existing ADS solution. SAE J3131’s reference architecture specifically calls out a “safety checker” as a key element to a safe-by-design ADS.[9] Formal Models like RSS can perform the safety checker function. The assurance provided by formal models like RSS allow for significant innovation within the ADS implementation, since new AI algorithms or other software updates can be introduced into the ADS

[9] https://www.sae.org/standards/content/j3131/.

without impacting safety as the formal model for safety remains unchanged.

Additionally, formal models for ADS planning are inherently flexible due to the parameterized nature of their design. As noted in Question 1, the regulator can dynamically adjust the desired level of ADS performance within the bounds of the chosen parameters if needed, without hampering innovation; nothing in the ADS needs to change – the ADS developers can make necessary enhancements without compromising safety. By simply changing the values for the parameters for reasonably foreseeable assumptions, adjustments can made to ADS behavior – just like changing the speed limit on a road does not require retraining of the driver but does affect their behavior.

Question 5. *How could the Agency best assess whether each manufacturer had adequately demonstrated the extent of its ADS' ability to meet each prioritized element of safety?*

Our proposed framework provides a straightforward way for the Agency to assess whether the manufacturer meets each prioritized element of safety. NHTSA can assess the real-world performance of the ADS against the defined requirements.

<u>Perception:</u>

In advance of deployment, vehicle manufacturers can provide test results on the accuracy of the perception system to show that appropriate validation data has been collected to prove the MTBF within the level defined by the regulator. Post deployment, ADS operators would only need to provide data on the hours of operation of the ADS and the number of failures that occurred within that timeframe to evaluate whether it indeed met the vehicle level failure rate (or MTBF) target. This would allow the Agency to continuously and efficiently assess manufacturer compliance without requiring large volumes of potentially proprietary data or large teams of experts to analyze and interpret the data. NHTSA's AV TEST Initiative infrastructure could easily be extended to support such data collection and reporting for commercially deployed vehicles.

To assess the ADS perception function, NHTSA could request evidence of offline (or online) test results along with information about the data set used by the manufacturer during testing (e.g., distribution of hours of testing conducted during daytime, nighttime, dry conditions versus wet conditions, etc.) in order to evaluate the results of the testing. Note again that testing of perception systems can be done efficiently offline, orders of magnitude faster than real-time driving, and benefit from a clear "ground truth" such that there is no subjectivity as to whether a perception system passed or failed a perception test.

Planning:

To assess the ADS planning function, NHTSA could evaluate the performance of the planning element within a simulator that has recreated known crash typology, such as NHTSA's pre-crash scenarios, or in the real world as part of a "driver's test" to assess whether the ADS is performing according to the reasonably foreseeable assumptions defined in a formal, risk-minimizing safety model. Finally, NHTSA could evaluate the formal model through review of formal verification evidence provided by the ADS developer.

The Institute for Automated Mobility in the state of Arizona is using existing traffic infrastructure to evaluate the performance of traditional vehicles against RSS-derived metrics[10], proving that metrics derived from formal models like RSS are not only technology-neutral (as all vehicles – even human driven – are being assessed) but that such an assessment can be done efficiently, often using existing infrastructure. NHTSA, through its consumer information program, NCAP, could also conduct performance tests of ADS-equipped vehicles just as it does for human driven vehicles.

Lastly, the Agency could assess compliance with the process measure requirements (see Question 1) of our proposed safety framework through safety self-assessment reports and/or independent audits to confirm

[10] https://saemobilus.sae.org/content/2020-01-1206.

compliance with the ADS process measure requirements, as is common today with traditional vehicles in the automotive community.

Question 6. *Do you agree or disagree with the core elements (i.e., "sensing," "perception," "planning" and "control") described in this notice? Please explain why.*

Intel agrees that the Agency has accurately described the core elements of an ADS. As explained in Question 2 above, we believe that of these elements, NHTSA should focus its monitoring and safety oversight efforts on perception and planning. The modality of sensing is not relevant to the accuracy of perception (i.e., sensing can be covered via perception), and planning – where decision making happens – is the place to ensure that the ADS does not have lapses in judgement and can navigate safely within the bounds of what has been defined by the regulator as reasonably foreseeable. Additionally, the control element is not necessary to include in a framework for assessing ADS safety, as control functions are already well covered by existing regulation and test procedures.

Question 7. *Can you suggest any other core element(s) that NHTSA should consider in developing a safety framework for ADS? Please provide the basis of your suggestion.*

First, NHTSA should ensure that commercially deployed ADS-equipped vehicles, especially SAE Level 3 vehicles, can sufficiently support situations that require an emergency response. It is important than an ADS be able to properly respond to safety critical events, including evasive maneuvers, at, or above, human-level capabilities. In this way, the ADS can support crash avoidance maneuvers "by design."

For example, consider an ADS-equipped vehicle traveling in the center lane in a multi-lane interstate highway. If the ADS-equipped vehicle suddenly encounters an obstacle (e.g. a disabled vehicle or object suddenly exposed by a last second lane change of the vehicle that the ADS was following), the ADS-equipped vehicle must be able to perform an emergency response to change lanes to avoid to this common crash situation. However, if the ADS was designed with only forward facing sensors, it would have no ability to perform an immediate response to evade a common crash scenario.

Consider further if this vehicle was designed to meet SAE Level 3 automation. Under SAE Level 3, the human driver must perform the fallback-ready user function. Studies have shown a wide variance in the amount of time it takes for a human fallback-ready user to take over the automation function, with a least a few seconds often being required from being given a take-over request.[11] In this described scenario, the vehicle's reliance on the human operator as the fallback-ready user would likely lead to frequent crashes in common events requiring an emergency response. The fallback-ready user cannot be expected to provide immediate time sensitive responses – the ADS must have that responsibility, even in a SAE Level 3 ADS. NHTSA should require that vehicles of any automation level (SAE Levels 3-5) be sufficiently capable of handling situations which could be reasonably expected within the designated ODD that would require an immediate response, at least to the level or above, of that of a human driver.

The Automated Vehicle Safety Consortium ("AVSC") recently published a best practice for describing the ODD[12] which states that, "Identification of operational constraints involves testing the ADS on the subject road/network route in order to identify objects, zones, conditions, events, etc., that the ADS-operated vehicle will be designed to avoid. This supports the concept of a system that is "safe by design" because the ADS operation is only allowed in areas that are understood and appropriate to its capabilities." Intel fully agrees with this this safe-by-design approach. A SAE Level 3 vehicle without the ability to change lanes in response to a sudden situation requiring an emergency response would certainly not meet this guidance.

[11] Zhang, Bo & de Winter, Joost & Varotto, Silvia Francesca & Happee, Riender & Martens, Marieke. (2019). *Determinants of take-over time from automated driving: A meta-analysis of 129 studies*.
[12] https://saemobilus.sae.org/content/avsc00002202004.

As the driver cannot be responsible for the immediate response of the vehicle, a critically important result of following this approach is that any SAE Level 3 vehicle that supports autonomy with an artificially limited set of capabilities within the ODD must be equally as capable as an automated vehicle that supports the full set of reasonably foreseeable behaviors within the ODD; otherwise, the vehicle will not be able to conduct the necessary emergency maneuvers.

Over the longer term, Intel encourages NHTSA to define test procedures and/or provide test frameworks so that there can be an apples-to-apples comparison of ADS-equipped vehicles from different manufacturers. This would reduce confusion generated by assertions of compliance with NHTSA's guidelines using proprietary methods not subject to independent review or audit. For example, the U.S. Department of Transportation's Virtual Open Innovation Collaborative Environment for Safety ("VOICES") program, which currently is tasked with creating a distributed synthetic test environment for cooperative automation, could be redirected to providing a shared test infrastructure for testing the planning element of an ADS against any pre-crash scenarios NHTSA would like to evaluate against. This would much better serve industry and provide an apples-to-apples mechanism to evaluate the planning performance of ADS-equipped vehicles.

Question 8. *At this early point in the development of ADS, how should NHTSA determine whether regulation is actually needed versus theoretically desirable? Can it be done effectively at this early stage and would it yield a safety outcome outweighing the associated risk of delaying or distorting paths of technological development in ways that might result in forgone safety benefits and/or increased costs?*

Intel's proposed approach provides a two-step path that enables effective regulations in the near term, and over the long-term facilitates the benefits of ADS without delaying or distorting technological development. The current U.S. framework for motor vehicle regulation is premised on the presence of a human driver operating the vehicle and is incompatible with the abilities and needs of an ADS. For ADS to be deployed at scale, the automotive industry will need regulatory certainty and clarity in a timely manner. Additionally, for the U.S. to position itself as a global leader in advanced automated vehicle safety technologies, proactive regulation for ADS will be necessary. We estimate that industry will be ready to deploy this technology starting in 2022; other leading countries (e.g., France, Germany, etc.) are readying regulatory frameworks to meet those timelines.

Performance-oriented metrics, based on setting a minimum vehicle level failure rate (e.g., the MTBF) along with clear definitions of the reasonably foreseeable assumptions that an ADS should make about other road users can indeed be set today, thereby providing critical guidance and important clarity for the industry on the minimum-level of safety performance required by government and society. Established, clear performance-based metrics are independent of any technology choice and will in fact accelerate the safety benefits and acceptance of ADS-equipped vehicles. If NHTSA does not prescribe clear performance metrics and does not define values for what is reasonably foreseeable for use in planning systems, this will stunt the industry's ability to deploy ADS-equipped vehicles that are acceptable to society and delay the safety benefits that can be realized from ADS.

Lastly, NHTSA's selection of those metrics and assumption values is not 'final'. The benefit of an ADS that incorporates a parameterized formal safety model for planning is that it can easily be updated to support different values for the reasonably foreseeable assumption parameters, meaning that any performance requirements issued by NHTSA can easily be adjusted over time if needed.

Question 9. *If NHTSA were to develop standards before an ADS-equipped vehicle or an ADS that the Agency could test is widely available, how could NHTSA validate the appropriateness of its standards? How would such a standard impact future ADS development and design? How would such standards be consistent with NHTSA's legal obligations?*

Our approach proposes a two-step process, consistent with NHTSA's statutory authority: near-term definition by NHTSA of performance-oriented metrics for ADS to complement development of FMVSS for ADS in the longer-term. As noted in Question 5 above, our framework provides a straightforward way for the Agency to assess each prioritized element of safety. NHTSA can assess the real world-performance of the ADS against the defined requirements for both the perception and planning functions utilizing existing test fleets (e.g., AV TEST Initiative) as well as simulation environments to assess the performance of the planning element against known crash scenarios.

Question 10. *Which safety standards would be considered the most effective as improving safety and consumer confidence and should therefore be given priority over other possible standards? What about other administrative mechanisms available to NHTSA?*

Safety standards (e.g., FMVSS) that require transparency in decision-making and through the use of open and transparent, formal safety models for planning (e.g. RSS) that are formally verified to maximally contribute to the safety goal are the best way to ensure safety and obtain consumer confidence. Planning systems with opaque safety arguments based on a promise to "trust me" that are centered on unexplainable proprietary AI based "predict the future" ADS planning systems should be discouraged. When industry and government experts come together to align on what parameters are reasonably foreseeable for ADS behavior as is being done in IEEE 2846 – and governments publicly set the values for those boundaries in advance of deployment – society can reap the benefits. Failure by regulators to set these boundaries will stunt industry growth and shift deployment to other geographies where governments are providing the necessary clarity.

Question 11. *What rule-based and statistical methodologies are best suited for assessing the extent to which an ADS meets the core functions of ADS safety performance? Please explain the basis for your answers. Rule-*

based assessment involves the definition of a comprehensive set of rules that define precisely what it means to function safely, and which vehicles can be empirically tested against. Statistical approaches track the performance of vehicles over millions of miles of real-world operation and calculate their probability of safe operation as an extrapolation of their observed frequency of safety violations. If there are other types of methodologies that would be suitable, please identify and discuss them. Please explain the basis for your answers.

Societal expectations of ADS-equipped vehicles are that they must be at least as good as human drivers. For planning functions, formal safety models, like RSS, which incorporate rule-based approaches provide transparency and confidence that an ADS will make the right decisions for the right reasons. Rule-based methods are open, transparent, technology-neutral and have already been proven in other industries such as avionics to deliver a very high level of safety. Furthermore, formal models can be verified to produce the correct output whether the vehicle has driven 1 or 100 billion miles.

Reliance on statistical approaches to track the planning performance of an ADS-equipped vehicle is a less effective method that would require tens of billions of tracked performance data in order to understand if the vehicle in fact meets defined performance requirements or if it had just gotten lucky. Extrapolating to billions of miles while observing only a fraction of that distance, is a less effective method to evaluate safety, as even the best drivers can go long stretches of driving without an incident.

It is also very easy to generate a significant number of miles by testing in less complex environments. If an ADS-equipped vehicle was tested in perfect weather conditions on roads with very few vehicles or pedestrians, is it better than an ADS-equipped vehicle that has been tested a tenth as many “miles” but in a complex urban environment in all times of day and in a wide variety of different weather conditions without a single incident? What about an ADS system that has been tested one trillion “miles” within a simulator, on synthetic roads with abnormal virtual pedestrians that have no relationship to any real world environment or real world human behavior? Not all “miles” are created equal, and a statistical argument for planning functions based on the number of miles driven (in the real world or in simulation), is extremely difficult to assess and understand the comparative performance to a human driver.

For ADS perception systems, statistical evidence of the MTBF is a suitable approach, as perception systems by the nature of the technology, are always probabilistic and such perception systems, unlike planning systems, can be efficiently validated offline with real world data.

Question 12. *What types and quanta of evidence would be necessary for reliable demonstrations of the level of performance achieved for the core elements of ADS safety performance?*

With our proposed framework, NHTSA need only to collect simple evidence of failures in the field and total hours of operation to verify whether the performance of the ADS meets the Agency's defined levels of risk. Mean Time Between Failure, expressed in hours of operation of the ADS, is a straightforward performance target to define and easy for industry to provide evidence of (i.e., submitting the total number of hours the ADS operated with the number of failures that occurred during that time).

If NHTSA seeks additional evidence related to the perception function, it could also evaluate the size, characteristics, and distribution of the validation data derived from the ADS perception system to ensure it meets the defined MTBF and appropriately reflects the intended ODD of the ADS deployment. As noted above, planning systems that utilize a formal safety model like RSS can be verified without the need for a quanta of evidence and can be efficiently verified within a simulation environment.

Question 13. *What types and amount of argumentation would be necessary for reliable and persuasive demonstrations of the level of performance achieved for the core functions of ADS safety performance?*

ADS developers can publish their performance data including the total hours of operation and total number of at-fault crashes for its ADS, which would provide meaningful and transparent evaluation of the performance of the ADS fleet. They can also publish the values of the reasonably foreseeable assumptions that they are using in their planning elements to ensure compliance with regulatory requirements to use specific values for specific road users and/or environments. Developers can publish the results of simulation tests of the planning element having demonstrated complete coverage for known crash types defined by the regulator.

Finally, for the core function of perception, ADS developers can publish evidence of their perception tests against validation data sets (or real world performance data), along with information about the characteristics of the perception data sets so that it is clear whether the validation data sets are sufficient and consistent with the intended deployment environment of the ADS.

Question About NHTSA Research

Question 14. *What additional research would best support the creation of a safety framework? In what sequence should the additional research be conducted and why? What tools are necessary to perform such research?*

NHTSA should prioritize derivation of values for reasonably foreseeable assumptions about the behaviors of other road users, consistent with the framework provided by IEEE 2846. Regulation of such values provides important clarity to industry on what is reasonably foreseeable for ADS performance and is consistent with regulatory approaches in other parts of the world, such as that of the UNECE. In addition, NHTSA data sets such as the National Automotive Sampling System (NASS), General Estimates System (GES), and pre-crash scenario typology studies provide a data-driven path to understanding and baselining current human driver performance in various scenarios.

Questions About Administrative Mechanisms

Question 15. *Discuss the administrative mechanisms described in this notice in terms of how well they meet the selection criteria*[13] *in this notice.*

Voluntary Mechanisms

Safety Assessment and Other Disclosure/Reporting

Disclosure/reporting such as the safety self-assessment and NHTSA's AV TEST initiative are most beneficial for assessing transparency. Reporting by ADS developers about their use of process and engineering measures to ensure safety during the design, manufacture, testing and deployment of ADS-equipped vehicles can help build public confidence and acceptance of ADS. Reporting can also enhance NHTSA's efficiency by allowing the Agency to gain insights about ADS design and how ADS developers have by-design minimized safety risks without having to perform costly and complex independent assessments. This efficiency can also help the Agency optimize its resource requirements. Additionally, reporting and disclosures can provide NHTSA with information and criteria that can inform the Agency in its development of recommendations and/or standards for consistent and reliable assurance of safety, allowing for technology neutrality, equity, and consistency with market innovation.

Considering the benefits that reporting and disclosures can provide to the Agency in the near and mid-term, we support NHTSA utilizing mandatory reporting/disclosures to collect ADS-related safety data and other information to inform the development of safety standards for ADS.

NCAP

NHTSA should opt to add an ADS competency evaluation into NCAP. This type of evaluation is already being conducted by EuroNCAP[14] to measure the relative performance of an ADS navigating a variable environment and complex set of interactions and stimulus. This approach could assist NHTSA in developing criteria to

[13] *See* criteria listed in NHTSA ANPRM: Consistent and Reliable Assurance of Safety; Technology Neutrality/Performance-Based; Predictability; Transparency; Efficiency; Equity; Consistent with Market-Based Innovation; and Resource Requirements.

[14] https://cdn.euroncap.com/media/58812/euro-ncap-ad-test-and-assessment-protocol-v10.pdf.

objectively assess a common standardized minimum level of safety and performance outcomes. An ADS competency evaluation in NCAP could be technology-neutral/performance-based, provide predictability and transparency to manufacturers and the public, and would be equitable for all manufacturers. NCAP can encourage safety advancement and swift adoption of performance improvements for ADS, consistent with market-based innovation.

Operational Guidance

Through operational guidance, NHTSA can begin to establish a consistent and reliable assurance of safety that ensures objective assessment of each manufacturer's methods to meet a defined minimum level of safety for ADS (e.g., the MTBF). The Agency's guidance can also express its desire to remain technology-neutral and focus on performance-oriented terms to provide manufacturers broad choices and flexibility to develop and introduce new technologies. Guidance can be provided efficiently as it can be modified and implemented more quickly, as well as maximize the Agency's resource requirements for safety oversight. Guidance does not unnecessarily inhibit innovation since it does not need to meet strict requirements of standards. Additionally, it can outline performance outcomes to demonstrate safety, thereby encouraging predictability and equity.

While Intel strongly prefers binding performance standards for ADS safety, the industry needs clearly established performance metrics for ADS safety and defined values from the regulator on reasonably foreseeable assumptions about the behavior of other road users in the near term to enable commercial deployment at scale in a timely manner. The lack of performance metrics and values for reasonably foreseeable assumptions will delay this important safety technology and result in continued confusion within the industry regarding performance requirements for ADS-equipped vehicles, allowing bad actors to put unsafe vehicles on the road due to the lack of clear guidance on minimum levels of safety. Performance data, along with Process and Engineering measures, provide meaningful information to develop metrics today and already prove the efficacy of ADS-equipped vehicles that are safe-by-design.

Regulatory Mechanisms

Mandatory Reporting/Disclosure

Requiring mandatory reporting/disclosures in the context of ADS deployment may be beneficial for the Agency to exercise oversight and safety monitoring of ADS operation. This includes information such as annual reporting of the MTBF information regarding all vehicles in operation.   Intel supports the use of mandatory reporting/disclosures to collect ADS-related safety data and other information that can inform the development of safety-performance standards. Under our framework, MTBF reporting would be useful data for the Agency to assess ADS performance.

FMVSS

Intel strongly supports the development of FMVSSs that define binding performance requirements for ADS safety. Regulating through FMVSSs will establish a consistent and reliable assurance of safety for ADS performance, eliminating the confusion that currently exists in the industry. A key benefit of federal safety standards is regulatory certainty and predictability, allowing manufacturers to anticipate the types of performance outcomes they will need to demonstrate the safety of their ADS. Also, FMVSS regulation is transparent, and grants equity to all manufacturers.

FMVSSs should be performance-based, and developed in a technology-neutral manner, consistent with market-based innovation. Our proposed framework is for NHTSA to prescribe performance-oriented metrics expeditiously to facilitate a high level of safety across the entire industry in the near-term while adopting FMVSS in the longer-term. Our framework focuses on the ADS perception and planning functions and includes industry-led process and engineering measures that can be verified through a transparent, technology-neutral formal safety model. Our approach is flexible and allows manufacturers to determine how best to meet defined parameters for reasonably foreseeable assumptions about other road user behaviors, and as pointed out above can be easily and efficiently tested.

Question 16. *Of the administrative mechanisms described in this notice, which single mechanism or combination of mechanisms would best enable the Agency to carry out is safety mission and why? If you believe that any of the mechanisms described in this notice should not be considered, please explain why.*

Intel prefers that NHTSA adopt a regulatory approach that establishes Federal safety standards for ADS while prescribing clear performance-oriented metrics in the near term. Safety standards would best enable the Agency to carry out its safety mission because they are binding for the entire industry and will provide regulatory certainty and predictability, thereby eliminating confusion about performance requirements for ADS-equipped vehicles. This certainty will create a consistent and reliable assurance of safety, facilitating the ability to deploy ADS-equipped vehicles at scale in the United States. We propose that the Agency ensures that its standards are technology-neutral and performance-oriented to give manufacturers flexibility to innovate to develop and introduce new technologies.

However, we understand that safety standards may take more time to implement. Therefore, the framework we propose is NHTSA should define and release clear performance metrics that align with this approach as a first step in the near-term. Under this approach, NHTSA should clearly define performance-oriented metrics for ADS safety (e.g., the MTBF) and incorporate industry driven process and engineering measures. Additionally, NHTSA should consider the use of formal safety models to assess the real-world performance of an ADS in a transparent and technology-neutral manner based on the philosophy of setting reasonably foreseeable assumptions about the behavior of other road users. Industry standard IEEE 2846 gives excellent guidance in this area, outlining a framework of parameters within which NHTSA can provide values for the reasonably foreseeable assumptions to industry for ADS performance.

Question 17. *Which mechanisms could be implemented in the near term or are the easiest and quickest to implement and why?*

The mechanisms that could be implemented in the near-term are operational guidance and mandatory reporting. Intel's preferred option is operational guidance that clearly establishes safety performance metrics that must be met by an ADS to enable commercial deployment of ADS-equipped vehicles at scale in a timely manner. The Agency's operational guidance should reflect the proposed safety framework and process and engineering measures outlined in our comments.

Mandatory reporting of at-fault failures in the field can verify whether the performance of the ADS meets NHTSA's defined levels of reasonably foreseeable risk. Mean Time Between Failure, expressed in hours of operation of the ADS, is a simple performance target to define and easy for industry to provide evidence of (i.e., total number of hours the ADS operated with the number of at-fault failures that occurred during that time). Mandatory reporting to NHTSA is already being conducted by manufacturers in a variety of contexts and can easily continue and be adapted for ADS safety.

Question 18. *Which mechanisms might not be implementable until the mid or long term but might be a logical next step to those mechanisms that could be implemented in the near term and why?*

Performance data, along with Process and Engineering measures provide meaningful information and already prove the efficacy of ADS-equipped vehicles that are safe-by-design. This information can be used to develop ADS safety; however, a variety of factors may hinder their implementation in the near term. Therefore, ADS-specific FMVSS may not be implementable for NHTSA until the mid or long-term. As described in our approach, Federal safety standards are the next logical step for the Agency following the establishment of performance-oriented metrics along with mandatory reporting/disclosure of ADS-related safety data and information.

Question 19. *What additional mechanisms should be considered and why?*

Any mechanism to prescribe transparent publication of performance metrics in the near-term would be beneficial in terms of enabling market development.

Question 20. *What are the pros and cons of incorporating the elements of the framework in new FMVSS or alternative compliance pathways?*

FMVSS provide regulatory certainty for manufacturers and ADS developers by defining ADS safety. Incorporating the performance-oriented metrics elements of our framework will establish a consistent baseline for the entire industry and reliable assurance of safety for the public and eliminate confusion surrounding performance requirements for ADS, thereby creating a pathway for ADS deployment at scale. Failure to expeditiously define performance metrics (e.g., the MTBF) and define requirements on what is reasonably foreseeable for an ADS to expect from other road users consistent with IEEE 2846 will stunt industry growth and shift ADS deployment to other geographies where governments are providing this necessary clarity, while allowing potentially unsafe vehicles on the road due to the lack of a minimum performance requirement. We prefer FMVSS over alternative compliance pathways over the long-term since Federal standards are binding and provide predictability for manufacturers and developers regarding the types of performance outcomes needed to demonstrate ADS safety. Standards also allow for an apples-to-apples comparison across the entire industry for ADS performance.

Question 21. *Should NHTSA consider an alternative regulatory path, with a parallel path for compliance verification testing, that could allow for flexible demonstrations of competence with respect to the core functions of ADS safety performance? If so, what are the pros and cons of such alternative regulatory path? What are the pros and cons of an alternative pathway that would allow a vehicle to comply with either applicable FMVSS or with novel demonstrations, or a combination of both, as is appropriate for the vehicle design and its intended operation? Under what authority could such an approach be developed?*

Intel supports flexible demonstrations of competence with respect to the core functions of ADS safety performance. As described in previous responses, our framework proposes NHTSA focuses on the perception and planning functions of the ADS. ADS perception can be measured through efficient offline tests or by observing real world events to determine if the vehicle meets the level of performance defined by the regulator. For the ADS planning function, Intel strongly recommends inclusion of a formal safety model, like RSS, which can be efficiently verified in simulation against known pre-crash scenarios. Statistical approaches, which extrapolate to billions of miles while observing (or simulating) only a fraction of the actual distance, are a less effective method to evaluate safety of planning functions and fail to differentiate between a statistically lucky test versus an ADS that is safe-by-design.

Additionally, the Agency could assess ADS competence through safety self-assessment reports and/or independent audits to confirm compliance with process measure requirements, as is common today with traditional vehicles in the automotive community.

These alternative pathways could be developed pursuant to NHTSA's authority under the Motor Vehicle Safety Act.

Questions About Statutory Authority

Question 22. *Discuss how each element of the framework would interact with NHTSA's rulemaking, enforcement, and other authority under the Vehicle Safety Act.*

Each element of our proposed safety assurance framework for ADS would be consistent with the current requirements under the Motor Vehicle Safety Act, and would not necessitate any changes to NHTSA's rulemaking, enforcement, and or other authority under the Motor Vehicle Safety Act. As required under the Motor Vehicle Safety Act, our approach is performance-oriented and the need for ADS-specific safety standards can be demonstrated through performance data from ADS-specific testing and data generated from a verified, formal safety model like RSS tested against known pre-crash scenarios. Such safety models are technology-neutral and allow the Agency to dynamically adjust the desired level of ADS performance if needed by simply changing the values of the reasonably foreseeable assumptions without stifling innovation. NHTSA can use simulation to verify the suitability of the values for the reasonably foreseeable assumptions for other road user behavior used within the safety model by testing against known or expected pre-crash scenarios. Current enforcement mechanisms can ensure that ADS-equipped vehicles that do not meet the Agency's performance standards are remedied in a timely and effective manner.

Question 23. *Discuss how each element of the framework would interact with Department of Transportation Rules concerning rulemaking, enforcement, and guidance.*

Our proposal to develop a framework for ADS safety incorporates the following elements: (1) Near-term performance-oriented metrics for ADS safety; (2) Defining a vehicle level failure rate for ADS performance; (3) Defining values for reasonably foreseeable assumptions in other road user behavior; and (4) Technology-neutral process and engineering measures, including a formal model for ADS planning. Each of these elements would be consistent with Department of Transportation Rules concerning rulemaking, enforcement and guidance, and would not require any modification to these requirements. As outlined in responses above, our framework is consistent with the Agency's authority and obligations under the Motor Vehicle Safety Act, NHTSA's authority to establish a performance-oriented safety framework for ADS-equipped vehicles, and its existing enforcement mechanisms.

Question 24. *If you believe that any of the administrative mechanisms described in this Notice falls outside the Agency's existing rulemaking or enforcement authority under the Vehicle Safety Act or Department of Transportation regulations, please explain the reasons for that belief.*

Intel believes that the administrative mechanisms described in this Notice and proposed within our safety framework are within the Agency's existing rulemaking or enforcement authority under the Motor Vehicle

Safety Act or Department of Transportation regulations.

Question 25. *If your comment supports the Agency taking actions that you believe may fall outside its existing rulemaking or enforcement authority, please explain your reasons for that belief and describe what additional authority might be needed.*

Intel’s comment submission does not support the Agency taking actions that fall outside its existing rulemaking or enforcement authority.

Appendix to Q1 Response

In the following Appendix, we are pleased to provide more details about formal safety models like RSS to provide more context regarding our proposed framework.

Formal Safety Models (e.g. RSS)

Fortunately, safety-by-design approaches, long proven in other industries like avionics[15], are being widely adopted in the ADS industry to facilitate planning decision-making and provide an ideal complement to probabilistic based perception functions.

RSS

Mobileye first published and contributed to industry its Responsibility-Sensitive Safety (RSS) model in 2017.[16] RSS is an open and transparent mathematical model for automated vehicle safety. RSS is technology-neutral and provides complete coverage for any driving scenario that an ADS may encounter within the bounds of reasonably foreseeable assumptions of other road user behaviors.

RSS is based on common-sense human notions of what it means to drive safely. RSS is open and transparent for all to see, and is based on general behavioral characteristics familiar to the best of human drivers:

1. Do not hit someone from behind
2. Do not cut in recklessly
3. Right-of-way is given, not taken
4. Be careful in areas of limited visibility
5. If you can avoid an accident without causing another, you must do it

[15] https://link.springer.com/chapter/10.1007/978-3-642-34281-3_2.

[16] Shalev-Schwartz, Shamma, Shashua. *On a Formal Model of Safe and Scalable Self-Driving Cars*. (2017). https://arxiv.org/abs/1708.06374.

Rules 1-2 define a longitudinal and lateral safety envelope which is the foundation of determining what constitutes a safe minimum following distance. These physics-based calculations account for the capabilities of the ADS while also incorporating reasonable assumptions on the behavior of other road users.

With clear definitions of what constitutes a minimum safe longitudinal or lateral distance, the ADS can then determine, in any driving scenario, on any road type, at any speed, whether it is currently at a safe distance with respect to other road users or if it must perform a *proper response* to restore the minimum safe distance. Formal proofs show that if an AV were to always perform a proper response to restore the minimum safe distance defined by RSS, then the AV should never be the cause of crash, and that if all other road users were to behave consistent with the model, then there would never be any crashes.

Rule 3 provides important clarity on what an ADS should do in cases where it may have the right-of-way but another road user is 'taking it' – e.g., another vehicle is running a red light when the AV has a green light. Rule 3 also enables the ADS to perform everyday merging behaviors that require assertive, but safe negotiation between other road users so that when necessary the AV can take the right-of-way as needed.

Rule 4 ensures that the ADS will consider the possibility of pedestrians or other road users being occluded behind parked cars or other obstructions.

Finally, Rule 5 covers the ADS's ability to take reasonable actions, if circumstances permit, in response to the unreasonable actions of other road users. In other words, if another road user behaves beyond the reasonably foreseeable assumptions, and the ADS can still avoid a crash without causing another, then Rule 5 states that the ADS must do so.

Formal safety models such as RSS are based on making reasonably foreseeable assumptions about the potential behavior of other road users. Just as humans make assumptions about the expected maximum deceleration rate of the vehicle they are following, or make assumptions about the possible maximum acceleration of a pedestrian, an ADS must also make these same assumptions in order to operate in the real world.

There are four key features of RSS.

- The first is that there are no contradictions. Namely, a proper response with respect to a specific vehicle will never contradict a proper response to another vehicle.
- The second is that RSS is efficiently verifiable. Safety models must be able to be implemented correctly. This is not all that obvious due to possible "butterfly effects" where a seemingly innocent action could lead to an accident in the longer future. We ensure this will not happen by following the inductive principle — a feature designed into RSS.
- The third property of RSS is the ability to cover all driving scenarios. We call it "completeness". By adopting the "worst-case under reasonable assumption" methodology, the RSS is provably complete.
- The aggregated result of those three principles leads us to the fourth property. If all agents follow RSS (and there are there are no other mechanical or infrastructure failures), then there should no accidents.

Support for an RSS-like approach is strong and growing:

- Following RSS's first publication in 2017, eleven leading automotive industry stakeholders including BMW, Daimler, Continental, Aptiv, FCA, VW and more, together endorsed the RSS approach via the "Safety First for Automated Driving" report[17], which was subsequently also approved as an ISO Technical Report, ISO TR 4804[18], and is now on its way to becoming an ISO Technical Standard.
- The RAND Corporation also highlighted RSS as the closest example of a leading measure of

[17] https://www.daimler.com/documents/innovation/other/safety-first-for-automated-driving.pdf.
[18] https://www.iso.org/standard/80363.html.

"roadmanship" in their comprehensive "Measuring Automated Vehicle Safety" report[19].

- SAE J3131[20] explicitly defines a "safety checker" element like RSS as a key Safety Layer of the ADS.
- ISO 21448's Informative Annex on how to apply the Safety of the Intended Function ("SOTIF") methodology to the creation of a safe driving policy provides RSS as an example consistent with the 21448 approach.

IEEE 2846

The culmination of RSS support in the industry is the creation of IEEE 2846, "Assumptions for Models in Safety-related Autonomous Vehicle Behavior". The IEEE 2846 Working Group led by Intel (Chair), Waymo (Vice-Chair), and Aurora (Secretary), consists of over 30 members including Mobility as a Service providers, leading automobile manufacturers, global auto industry Tier-1 suppliers, silicon providers, academic institutions and government entities, are collaborating on a consensus-driven standard that will define normative requirements about what reasonably foreseeable assumptions shall be considered by safety-related models in an ADS. Consistent with our commitment that the safety of AVs should not be proprietary, Intel contributed the RSS safety model to the IEEE 2846 Working Group for inclusion in future consensus-driven industry safety standards.

While IEEE 2846 supports only a subset of the capabilities of the RSS model, IEEE 2846 does provide a framework within which NHTSA can provide values for the parameters for reasonably foreseeable assumptions about the behavior of other road users so that industry can have clarity on the performance expectations of the ADS. So long as all other road users behave within the reasonably foreseeable assumptions, deterministic safety models like RSS enable ADS to make decisions that will positively contribute to the safety goal (e.g., do

[19] https://www.rand.org/content/dam/rand/pubs/research_reports/RR2600/RR2662/RAND_RR2662.pdf.
[20] https://www.sae.org/standards/content/j3131/.

not crash).

Technology-Neutral

Formal model approaches to safe ADS decision making are technology-neutral and provide critical transparency of the decision-making logic of an ADS, a criterion that NHTSA has highlighted is important for creating public trust. The use of unexplainable, proprietary AI approaches to safety that rely on the ability to accurately predict the future cannot offer the necessary data or information that regulators and society will need to understand and test the performance claims of the ADS, let alone enable stakeholders to figure out what happened when a collision occurs involving an ADS. Such "just trust us" approaches, will not engender essential public trust for AVs and could lead to a consumer backlash against this promising technology. Formal models like RSS establish trust through transparency and formal verifications of correctness to achieve safety-by-design, not safety-by-trial-and-error.

Formal models like RSS fit perfectly with NHTSA's suggested FMVSS approach to "require vehicles to be programmed to drive defensively in a risk-minimizing manner in any scenario within their ODD", and standards like IEEE 2846 provide NHTSA a technology-neutral framework to define what is reasonably foreseeable for other road user behavior, and to define the level of defensive risk-minimizing performance desired in specific driving scenarios.

Simulation/Testing

Simulation can be used to verify the correctness of the implementation of the safety model, and to study the effects of using different parameters in the model to assist in identifying the values that represent defined risk. Mobileye has released an open source implementation of RSS[21] that has been integrated into the open-source

[21] https://github.com/intel/ad-rss-lib.

CARLA[22] driving simulator. NHTSA pre-crash scenarios have been added to CARLA, so it is already possible to perform verification tests of an RSS implementation by assessing the test coverage of pre-crash scenarios in a simulated environment when using different values for the parameters that represent reasonably foreseeable behavior.

By clearly defining the boundaries and expectations of safety using a risk-minimizing approach like RSS, testing metrics can easily be derived from RSS to test vehicle performance in a technology-neutral manner. The Arizona Institute for Automated Mobility ("IAM") publication, "Driving Safety Performance Assessment Metrics for ADS-Equipped Vehicles",[23] defined a RSS-derived safety envelope and *proper response* based metrics to evaluate the performance of any vehicle (human or machine driven) operating in the real world against the reasonably foreseeable assumptions. In further research, the IAM is using existing roadway cameras and infrastructure to demonstrate how any vehicle can be evaluated against the proposed RSS-based metrics. What this research proves is that RSS-based metrics are not only technology-neutral, but that such metrics can provide quantitative criteria for the assessment of the performance of an ADS in the real world.

Conclusion re: Formal Safety Models

As a result, Intel strongly recommends that NHTSA consider the use of formal safety models like RSS within the planning element as this is the best way to provide certainty through safety-by-design and provide transparency to the public on what safety means for a machine driver. Intel also strongly recommends that NHTSA define specific values for reasonably foreseeable assumptions about the behavior of other road users to ensure a common level of safety by the AV industry.

[22] https://carla.org.

[23] https://saemobilus.sae.org/content/2020-01-1206.