Renault Group

The multidimensional complexity of Safety Assurance applied to Computer Vision for ADS & AD

Javier Ibanez-Guzman CEng PhD

23 October 2022

CONTENTS

RG

- **Context**
- **ADS & AV Architecture**
- **Added Complication: AI**
- **Complexity-Constraints:**
  - Pedestrian Detection,
  - Driving Monitoring
- **Multiple Approaches to Develop Safe Assurance: Trust**
- **Conclusions**

Context

RG

- Passenger Vehicles. Software has become the predominant factor: multiple processors, different data channels, interdependent functions, multiple sources of uncertainty, etc.
- Increasingly vehicle manoeuvres are being controlled by sensor-driven systems rather than drivers.
- Currently Autonomous Vehicles are being experimented in limited areas. Scaling-up has been found to be difficult.

- Premium vehicles → general vehicles
- Extended geographical areas and uses

- Robo-Taxi
- Shuttles
- Yard Logistics
- Mining
- …

Sensors
Computing
Connectivity
Artificial Intelligence

Actuating ADAS
- AEB Pedestrian
- Lane Keeping Assist
- Intelligent Driving Assistance
- Contextual ACC…

Autonomous Vehicles L4
- ODD: Restricted → Expanded
- Supervision: Human → Unmanned

Validation Complexity
System Complexity
Safety & Liability

CONTEXT

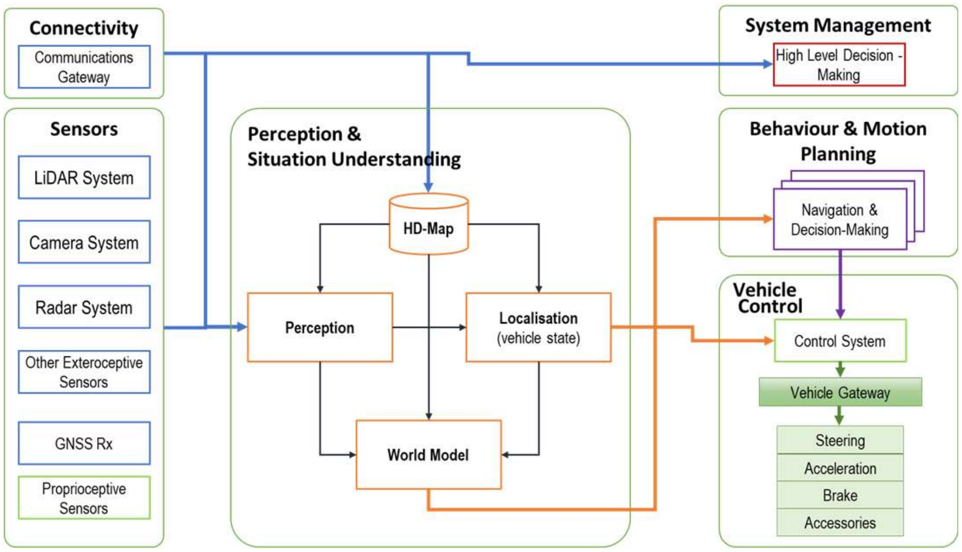
RG

- How **safe are current passenger cars**?
  - 1 fatality for an average of 81 Millions Km travelled (within the EU)
    - ✓ Say at an average speed of 50 km/h, 1.62 Million hours of usage
  - Issue: Reduction on accidents has flatten and in some cases, there is an increase
  - Waymo’s vehicles drove 32 million Km on public roads (2020). 2.3 Million Km in 2021.
  - Their autonomous cars have driven tens of billions of miles through computer simulations

- **Modern Passenger Vehicles**
  - ADAS: Sensor dependent functions e.g., AEB Systems, ACC, LKAS, ISA
  - Vehicles developed for safe operation over large geographic areas e.g., ISA (EU + +)

- **Autonomous Vehicles**
  - Currently all vehicles are **prototypes operating within confined ODDs** (NHTSA, 2021)
  - “Automated operations will only be feasible during the coming years within narrowly defined conditions” (Scientific American, 2021)

Typical Functional Architecture of an Intelligent Vehicle v1.1

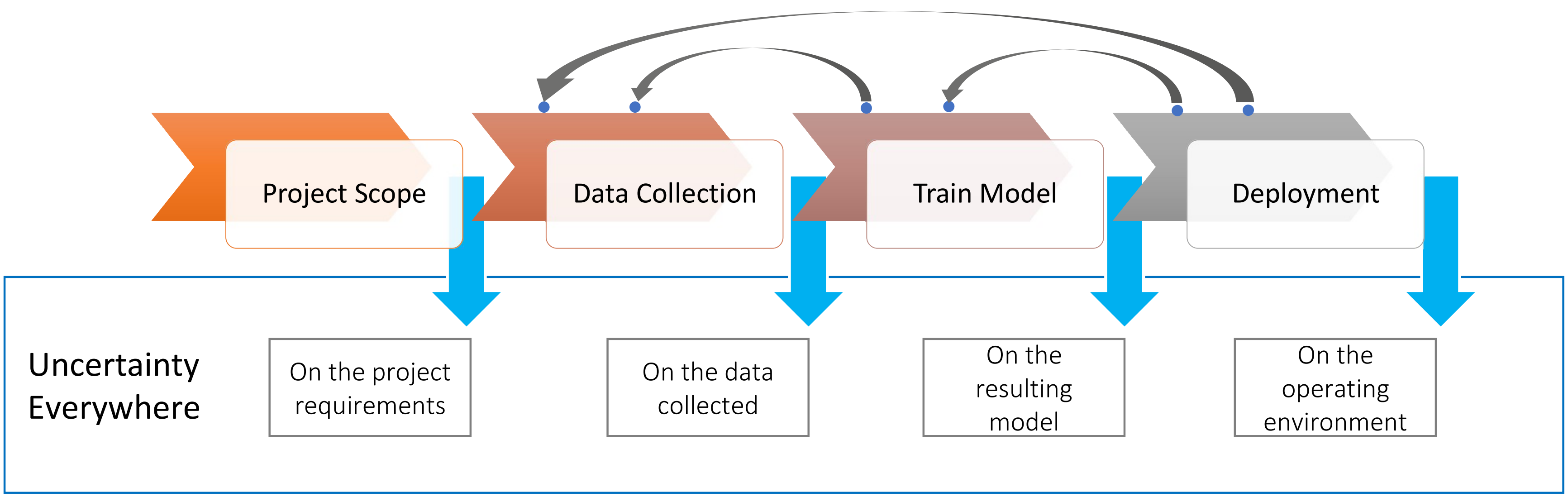


An important thread is the use of end-to-end approaches, that is the input will be data from the sensors and the output will be commands to the vehicle actuators.

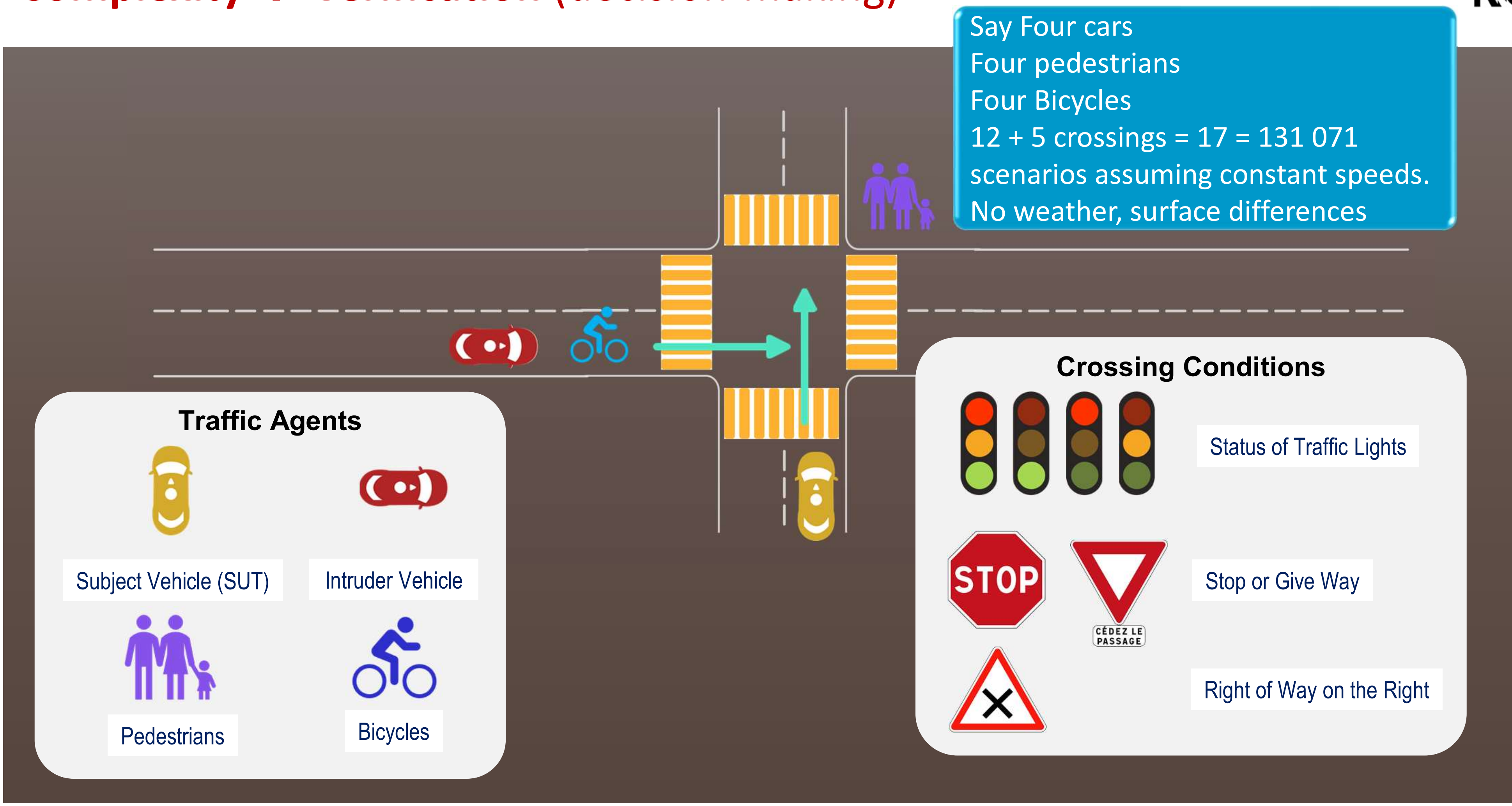
- ML distributed across the system.
  - Perception. Used extensively, the dominating force including multi-sensor fusion
  - HD-maps. Data extraction uses ML-based perception systems (off-line)..
  - Localisation. Whilst not the dominant approach, hybrid systems using perception systems or map-aided solutions also use ML.
  - World Model. Digital representation of immediate environment, ML prediction methods.
  - Navigation & Decision-Making. ML is not the dominant force,
  - Control System. Preferred approaches model-based methods. ML plays a secondary role.
  - High-Level Decision Making. ML is used to a lesser extent..

Intelligent Vehicles (Perception) : Data and more Data

Was the training data representative? What about the edge cases? Once we deploy, and need to update due to unforeseen Operational Design conditions, Update? What about the validation?



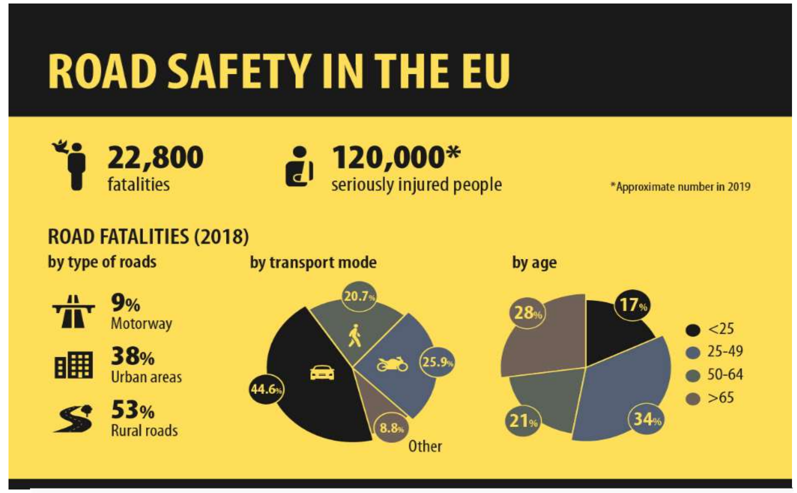
Complexity → Verification (decision-making)

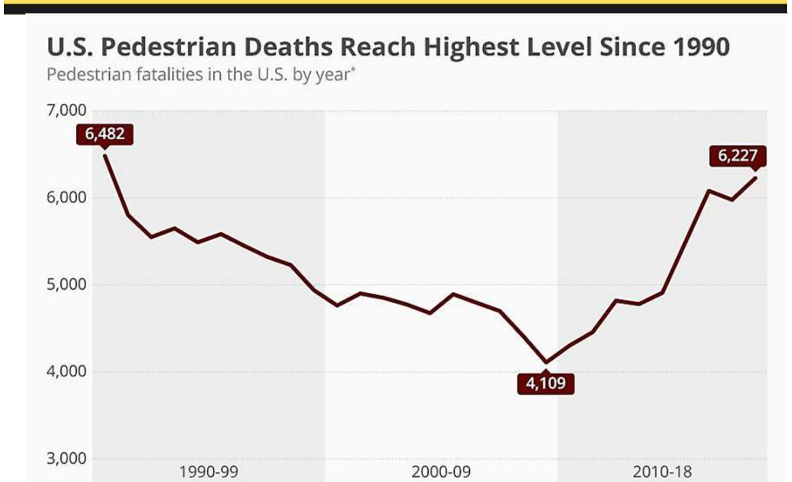


J.Ibanez-Guzman

UC1. Pedestrian Detection using a monocular camera, application to ADAS or AD systems.

- ADAS actuation is becoming an integral part of new vehicles. For example, the Autonomous Emergency Braking (AEB) system for pedestrians has been included in standard safety ratings since 2014.
- These systems depend on the ability to detect pedestrians in time. The algorithms used are all based on ML models.
- The determination of false positives can be done in field trials, the detection of false negatives is a major challenge.

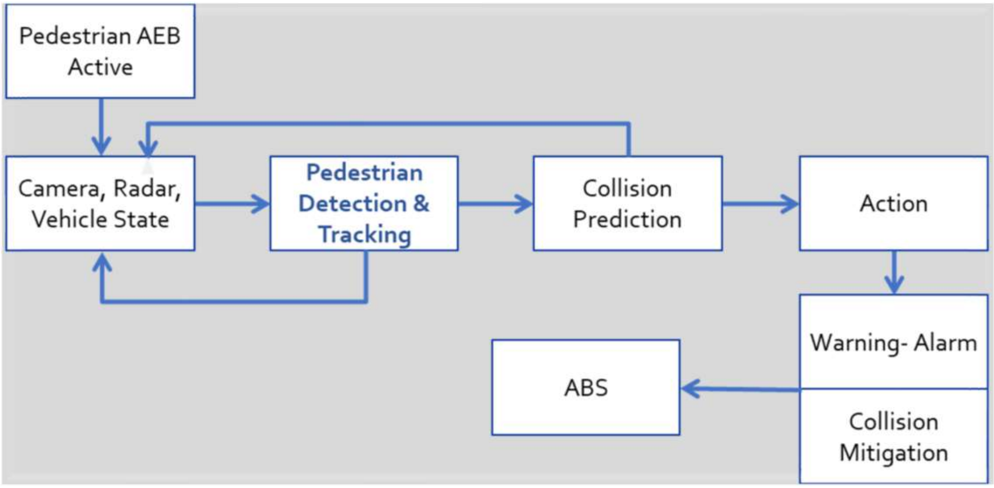




UC1. Pedestrian Detection using a monocular camera, application to ADAS or AD systems.

RG

- ADAS actuation is becoming an integral part of new vehicles. For example, the Autonomous Emergency Braking (AEB) system for pedestrians has been included in standard safety ratings since 2014.
- These systems depend on the ability to detect pedestrians in time. The algorithms used are all based on ML models.
- The determination of false positives can be done in field trials, the detection of false negatives is a major challenge.



Perception: Multidimensional Problem

FALSE NEGATIVE

System should have reacted but did not

- ✓ **Current False Positives and False Negatives: performance do not meet driverless safety goals**
  - → GAME(*)/10 = 10^{-8}/h fatal accident (highway)
  - → safety drivers
- ✓ **Current industrial perception capabilities do not meet AD L4 requirements**
- ✓ **Machine perception needs to understand robustly the driving scene - still a research subject**

(*) GAME = globalement au moins équivalent (at least as good as a driver)

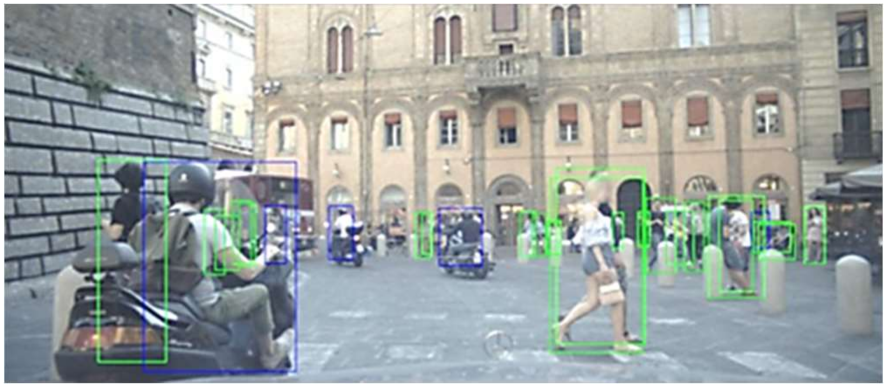
Perception: Multidimensional Problem → Verification of existing systems

RG

Pedestrian Detection (Automated Braking System): Fundamental as Safety Feature

"AI standard" to validate a Camera based Pedestrian Detection System

- Provide Statistical evidence to demonstrate whether their performance is satisfactory. Generally, done with respect to a ground truth which facilitates the quantitative evaluation. (Empirical Methods
- Concern: The Ground truth, for large amounts of images, is also generated using machine learning methods.
- To get a trustworthy assessment, the ground truth must be reliable. How can we demonstrate this in a credible manner?



Our Approach: Narrow the problem

Accidentology Data

- Road Network: streets, pedestrian crossings
- Areas: Urban, Peri-urban
- Day Conditions: daytime, normal weather
- Victims: adults, high proportion of elderly
- Scenarios: detailed classification

ODD Characterisation

System Ccapabilities

System Under Test Perception Capabilities

ML-Peerception System

- Pedestrian Classification
- Operation during daytime
- No age classification
- No depth information
- Needs cear visibility
- …

Examination of SUT capabilities wrt the ODD

Feasible ODD: specification

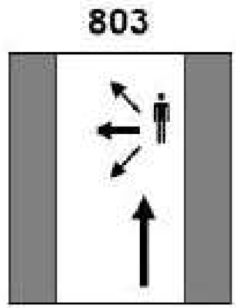
- Road Network
- Environment Conditions
- Geospatial constraints
- Other Traffic agents
- Target traffic agents: distribution

Applicable ODD

Specifications: Datasets, scenarios, tests,…

Perception: Multidimensional Problem

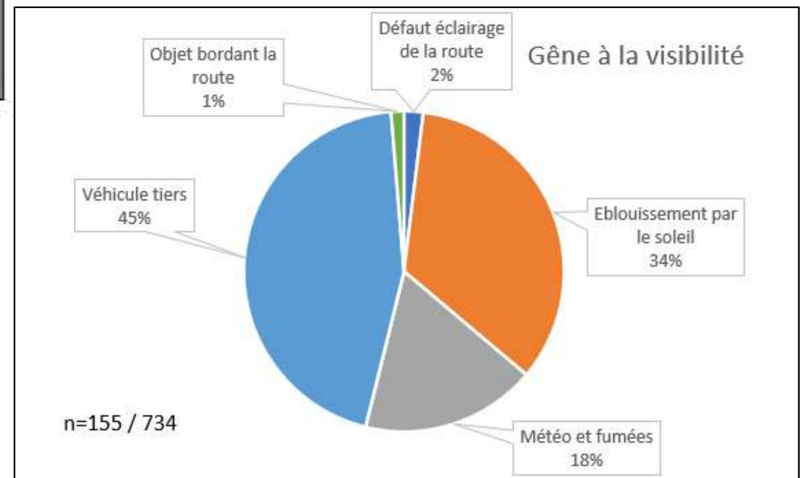
Pedestrian Detection (ABS): Approach

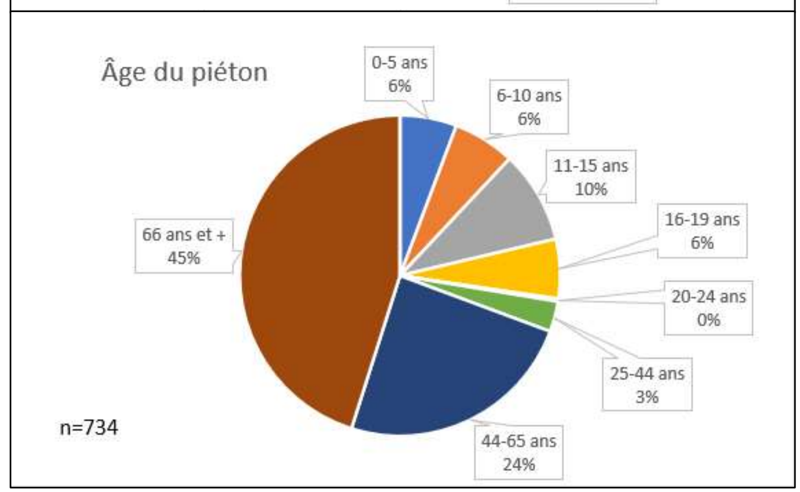


- Encapsulate the problem through the definition of the Operational Design Domain (oDD):
  - Conditions and geospatial descriptions

- ODD: Derived from Acidentology statistical data in France.
  - Traffic Scenarios: pedestrian crossing, intersection or jaywalking
  - Time of day: day / night /
  - Type of environment: urban, rural, highway ...
  - Weather conditions: low sun, rain, fog, etc.
  - Age group and "orientation" of injured pedestrians
  - Occlusion / masking
- ...

- Additional criteria
  - Max pedestrian detection distance depending on vehicle speed
  - False Negatives and False Positives Goals





UC2. Driver Monitoring and Assistance Systems for ADAS or AD systems.

- A significant portion of road accidents is attributed to drivers' inattention and aggressive behavior
- Driver comfort. Aim to provide a relaxed and comfortable driving experience.
- These systems depend on the ability to detect the face gestures (eyelids), gaze, etc.
- **Multiple Constraints: privacy, sensor location, non-invasive, etc**.
- Systems dépend on the use of cameras , ML-based

UC2. Driver Monitoring and Assistance Systems for ADAS or AD systems.: What to Observe?

RG

Symptoms Associated with Fatigue

- Frequent yawning
- Increased eye-blinking frequency
- Hard to keep eyes open
- Increased response time
- Nodding off

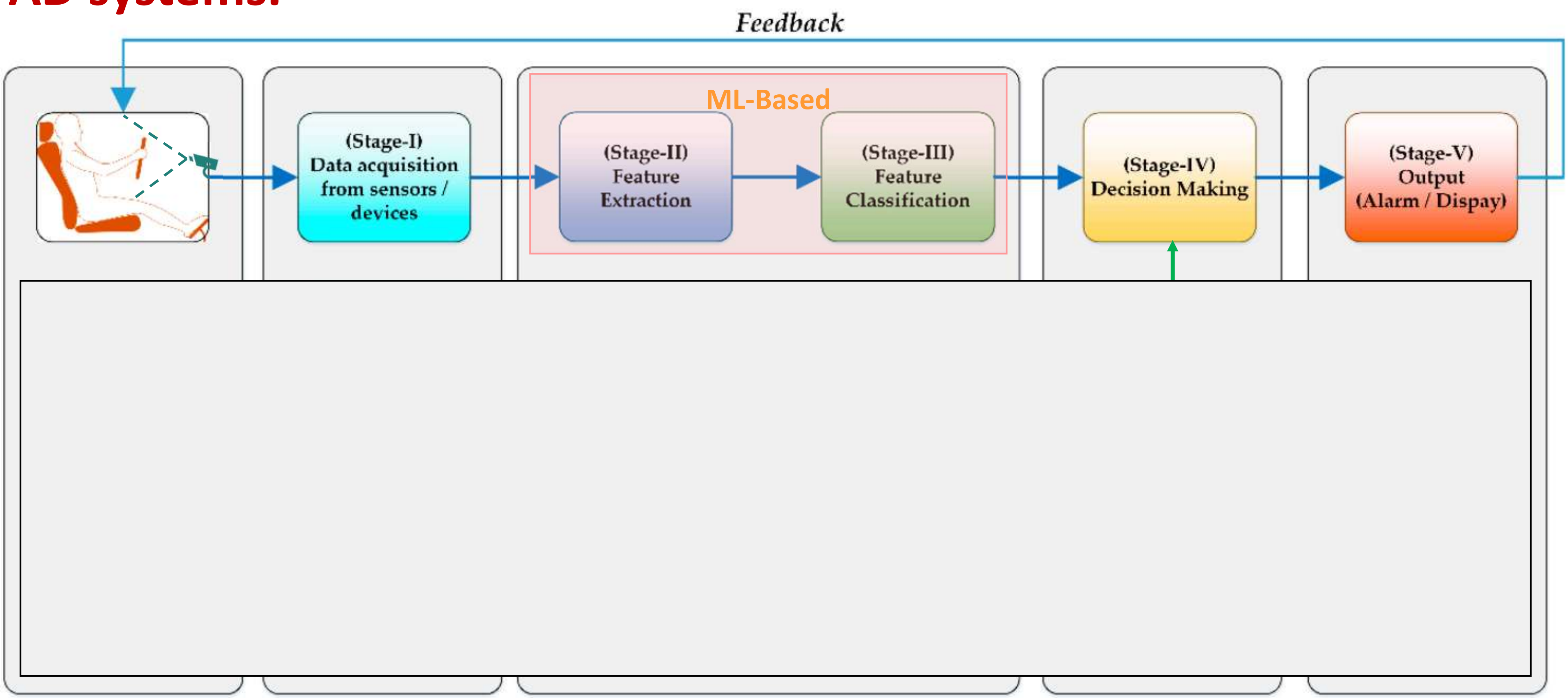
Driving Style

- Ignoring speed limits
- Opposite side driving
- Driving between two lanes
- Not using indicators whilst turning
- Overlooking road conditions

Distracted Driving

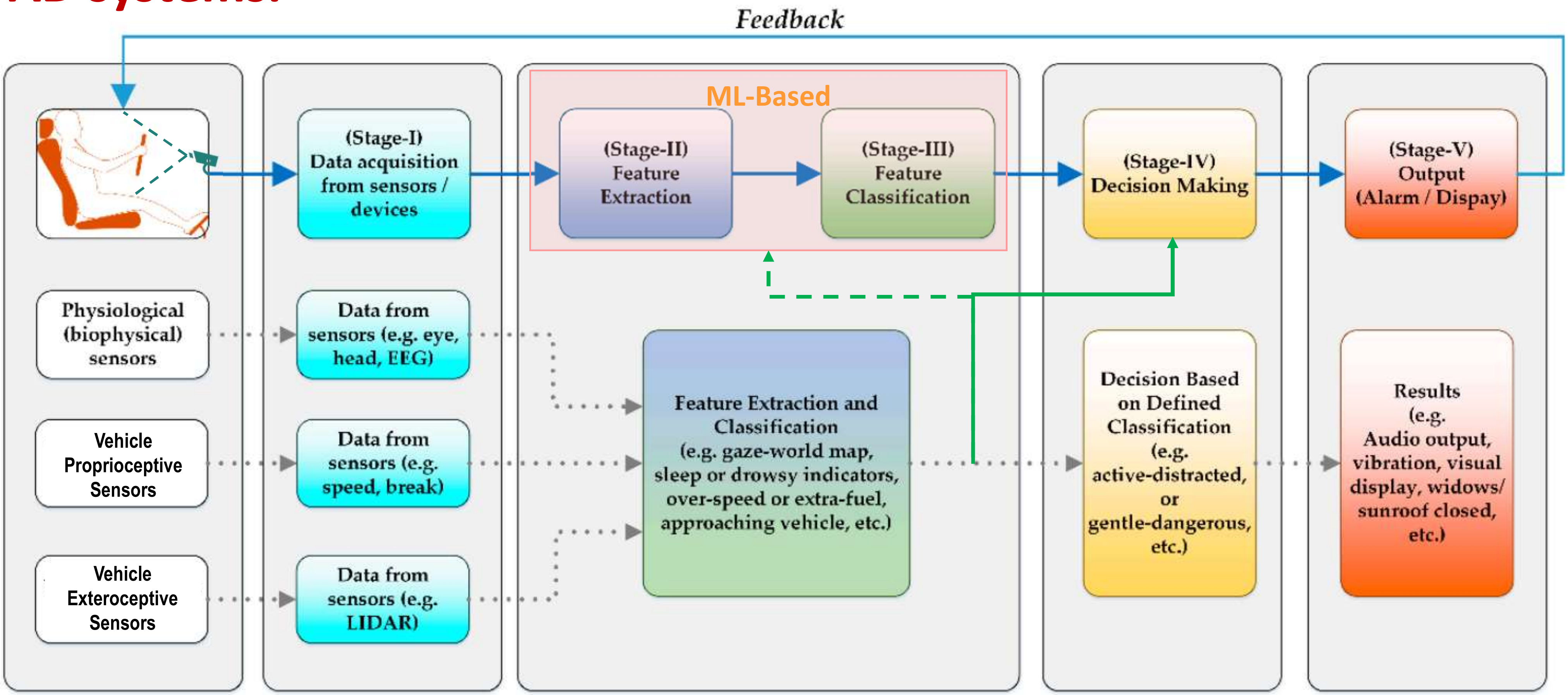
- Visual Distraction: driver's eyes off the road (gaze).
- Cognitive Distraction: driver's attention diverted away from driving
- Manual Distraction: Taking their hands off the wheel to perform another task.

UC2. Driver Monitoring and Assistance Systems for ADAS or AD systems.



Khan, M. Q., & Lee, S. (2019). A comprehensive survey of driving monitoring and assistance systems. Sensors, 19(11), 2574.

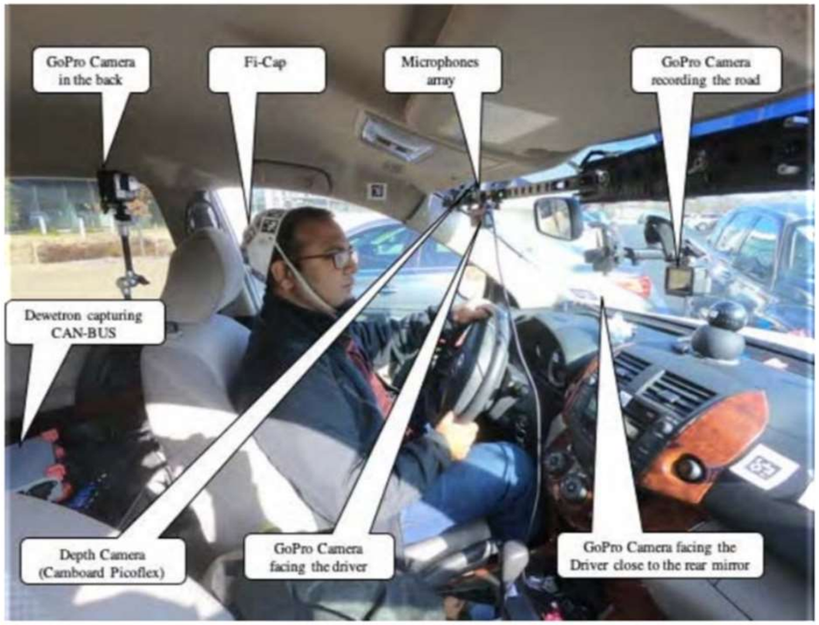
UC2. Driver Monitoring and Assistance Systems for ADAS or AD systems.



Khan, M. Q., & Lee, S. (2019). A comprehensive survey of driving monitoring and assistance systems. Sensors, 19(11), 2574.

UC2. Driver Monitoring and Assistance Systems for ADAS or AD systems.: ML-Models

RG



DATASETS

- DR(EYE)VE, day/night, diff. weather, urban, countryside
- DAD, naturalistic driving, urban, sub urban
- MDM, parked, daytime

Jha, Sumit, et al. "The multimodal driver monitoring database: A naturalistic corpus to study driver attention." IEEE Transactions on Intelligent Transportation Systems (2021).

Synthetic Datasets

- Specialist datasets (as above)
- DMD simulated and driving scenarios

Yudkin, P., et al., (2022). Hands-Up: Leveraging Synthetic Data for Hands-On-Wheel Detection. *arXiv preprint arXiv:2206.00148*.

UC2. Driver Monitoring and Assistance Systems for ADAS or AD systems.: ML-Models

RG

Context:

- DMS are becoming an integral part of upcoming driving assistance systems
- As vehicles are automated, knowing the driver state is a critical safety issue to warrant the takeover phase.
- Development based on datasets and different hybrid simulation setups…

Verification & Validation:

- How to do it? Datasets, Proving grounds, Driving Simulators, …
- We need real conditions, the observed subject shall show the appearances….

DATASETS

- DR(EYE)VE, day/night, diff. weather, urban, countryside
- DAD, naturalistic driving, urban, sub urban
- MDM, parked, daytime

Jha, Sumit, et al. "The multimodal driver monitoring database: A naturalistic corpus to study driver attention." IEEE Transactions on Intelligent Transportation Systems (2021).

Synthetic Datasets

- Specialist datasets (as above)
- DMD simulated and driving scenarios

Yudkin, P., et al., (2022). Hands-Up: Leveraging Synthetic Data for Hands-On-Wheel Detection. *arXiv preprint arXiv:2206.00148*.

Automated Driving System (ADS), and Safety Assurance

- **Safety assurance is the way you demonstrate that your Safety Management System (SMS) works.**
  - Systematic and ongoing monitoring and recording of your safety performance,
  - Evaluating your safety management processes and practices
  - Associate to modules and systems of the ADS system architecture and beyond.
  - A System is unsafe until proven otherwise
- **To gain Acceptability we need ADS Assurance Credible**
  - Engagement of stakeholders in defining safety goals and metrics
  - A SMS considering the entire lifecycle, demonstrating a organisational safety culture.
  - A Safety Case approved by independent safety experts
  - Provide evidence in an understandable manner addressing different stakeholders.
- **For highly ADS**
  - Driver fallback is reduced or unfeasible. The ADS needs to be self-reliant, self-resilient
  - The increased use of data driven solutions adds a higher level of complexity.
  - Deployments in multiple areas to optimise solutions results in growing challenges.

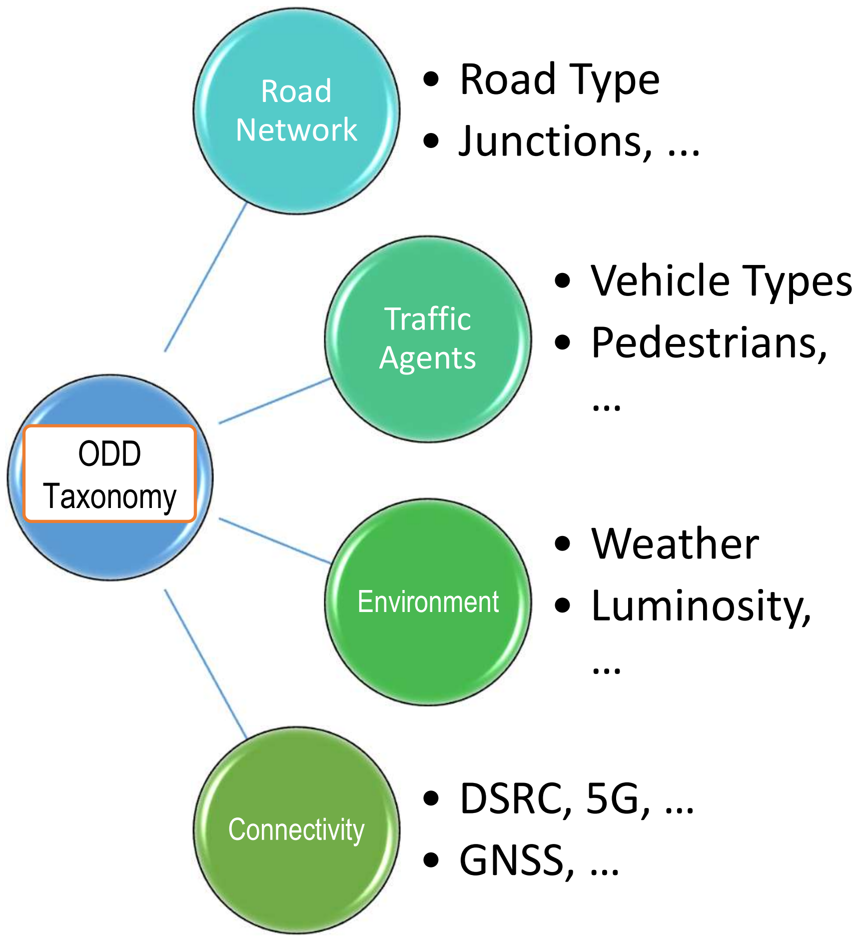
Multiple questions need to be addressed

Autonomous Stack → Behaviour Generation → for Safety Assurance Verification matters.

- What is the desired behaviour?
  - How can we expressed, so we can measure it?
  - E.g. How can we measure a safe behaviour?
- How to describe the desired behaviour?
  - Do we define some rule books?
  - Is there a hierarchy?
- What can we guarantee in an open ODD?
  - Nothing? Anyone can crash into you?
  - Can we demonstrate using theoretical methods? E.g. Formal Methods. Are they scalable?
- Can we quantify the AV level of residual risk?
  - What type of evidence can we generate?
  - What will be the position of the insurance company to determine the premium?

Verification: Isolating the problems: Operational Design Domain

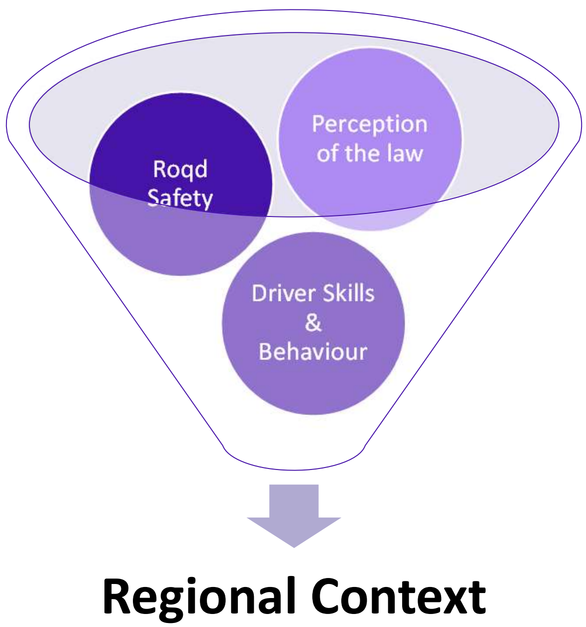
- **ODD** defines the workspace and conditions under which the AV is expected to safely operate.
- Emerged as substitute of Use Cases, in the automotive domain
- **Tenet:** Demonstrate the safety and operability of the system within a defined workspace and conditions.
- **Can we define the ODD through properties? Build descriptors so their coherence can be demonstrated?**
- This can be used in the AV to infer if it is operating within its ODD. Warrant its operation.



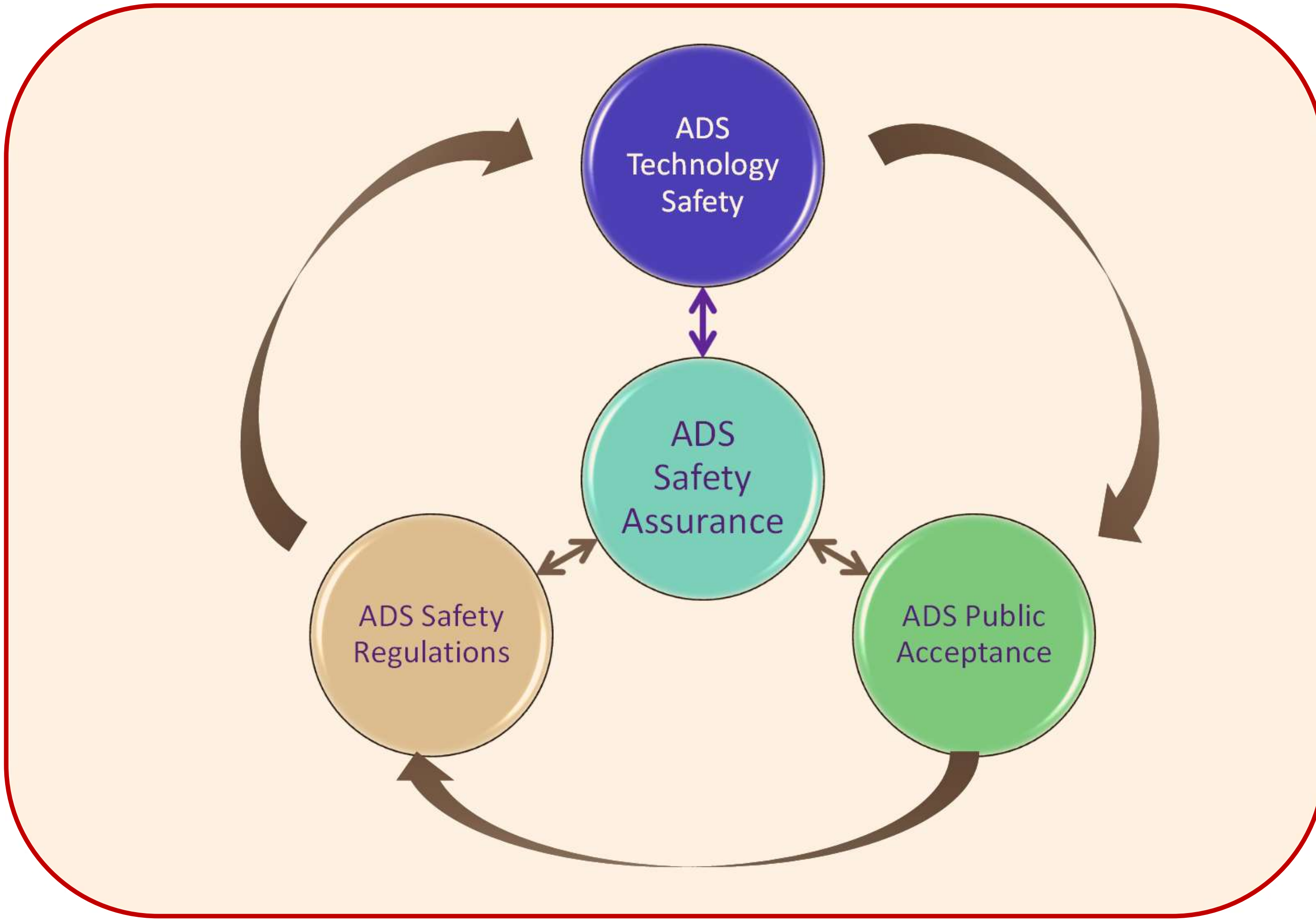
Safety Assurance and Safety Argument:
A multidimensional *DATA dependent* problem

DATASETS availability, scenarios , covering the expected ODD

- **Outlier/Anomaly Detection**. The detection of OoD (out of distribution Detection) inputs can be tackled by post-processing techniques that adjust the estimated confidence
- **Specific Domains/ Domain generalisation**. Diversity on the deployment of the application
- **Data Augmentation**. Optimize available data and increasing its amount, curating a dataset that represents a wide variety of possible inputs
- **Behaviour on Corner Cases**: Ensuring that AI-based applications behave correctly and predictably even in unexpected or rare situations
- **Adversarial Attacks**. Adversarial attack and aims at fooling an underlying DNN.

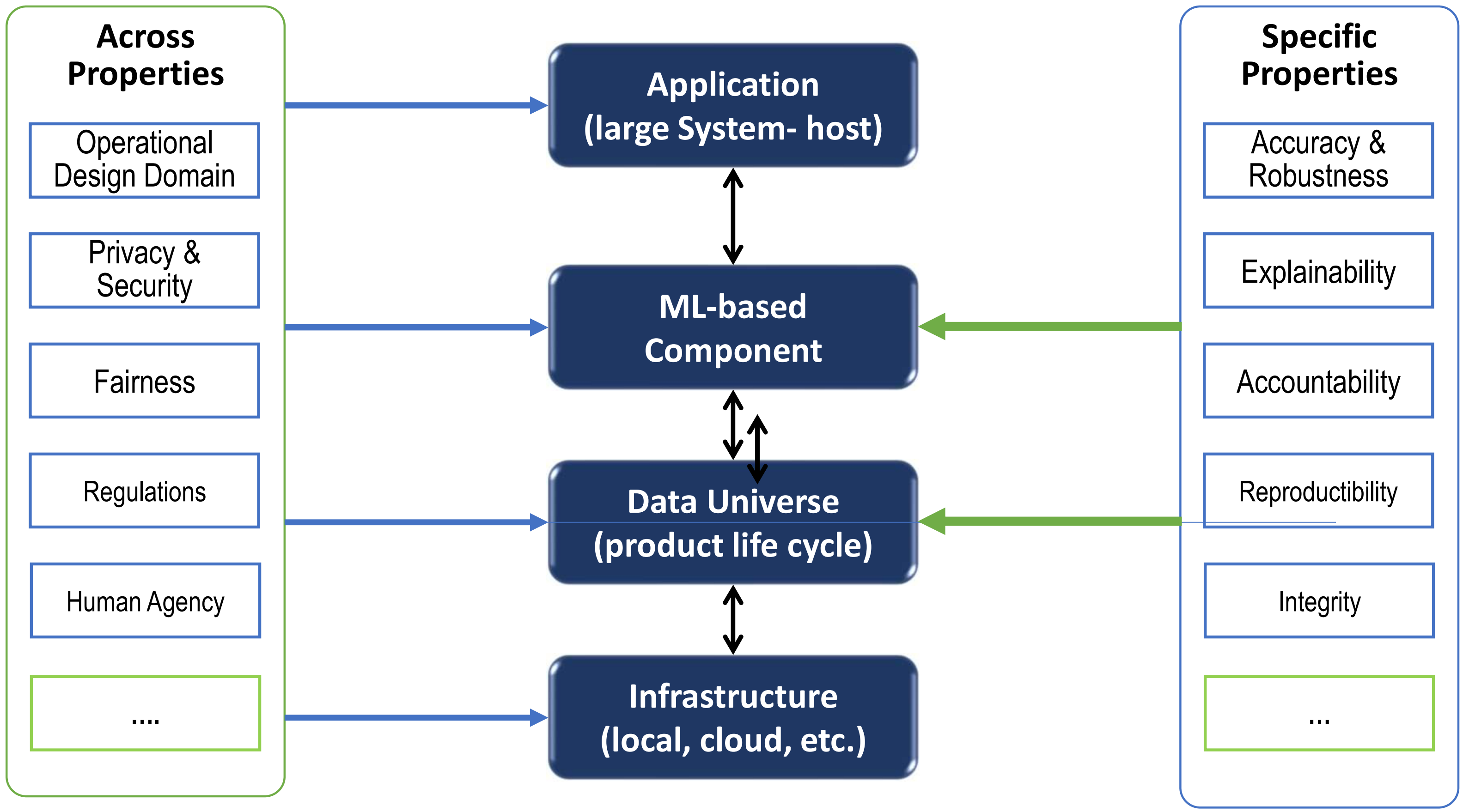


Safety Awareness: A holistic approach



Regional Specificities

A framework for Safety Assured ML-based Components



Kaur, Det al., (2021). Requirements for Trustworthy Artificial Intelligence – A Review. In L. Barolli, K. et al., (Eds.), Advances in Networked-Based Information Systems, Springer International Publishing.
Thuraisingham, B. (2022). Trustworthy Machine Learning. IEEE Intelligent Systems, 37(1)

Discussions:

- *Autonomous Vehicles represent perhaps the hardest challenge for AI:*
- Uncertainty : perception, localization, maps and the behaviour of traffic agents add complexity.
- One of the hardest challenges is to demonstrate that the systems are safe.
- V&V ADAS functions is more complex that what we envisioned.
- Much work is in progress on:
  - The use of systems engineering to make complex system “safe by design”
  - The extensive use of the ODD concepts to bound the problem
  - Introduction of formal methods in isolation
  - Methods for simulating ADS safety critical scenarios . Validation to an acceptable level of fidelity
  - Understand and explain ADS safety case findings to different stakeholders