Safety Assurance of Automated Driving Systems

Krzysztof Czarnecki
Waterloo Intelligent Systems Engineering Lab
Electrical and Computer Engineering Department

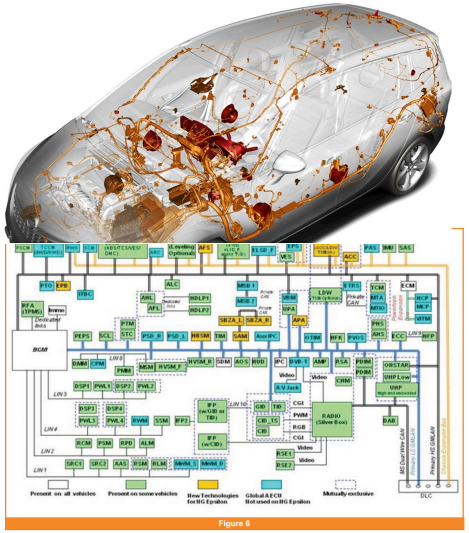
Some Background

2011-2015

Product-line Engineering of **Full Authority Digital Engine Control** (FADEC)

with Pratt & Whitney Canada

2014-2017



Design Exploration of **Automotive E/E Architectures**

with General Motors R&D

2016-2018

Full-stack Automated Driving Software (Autonomoose)

With Renesas, AAA

Operational Design Domain (ODD)

SAE J3016 Levels of Driving Automation

A set of **conditions** under which
the driving automation can operate a vehicle

Time of day
day
night

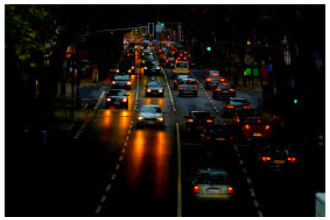
Types of roads
residential
urban
highway

Geographic area

Traffic conditions
stop-and-go
free flowing

Weather conditions
clear
raining
snowing
icy



Dynamic Driving Task (DDT) Fallback

Who performs the DDT
in the case of **system malfunction** or
when **leaving the ODD**?

Driver

Computer

Automated Driving Systems (ADS)

	3	4	5
SAE J3016 Levels of Automation	Conditional Automation	High Automation	Full Automation
Example	ADS for stop-and-go	Shuttle in geo fenced area	Robo Taxi anywhere
ODD	**limited**		**unlimited**
Fallback	Driver	ADS	

ADS Hazard Sources

Mature best practices

- Mechanical faults
- Electrical faults

ISO 26262

- Computer HW faults
- Computer SW faults

01100
10110
11110

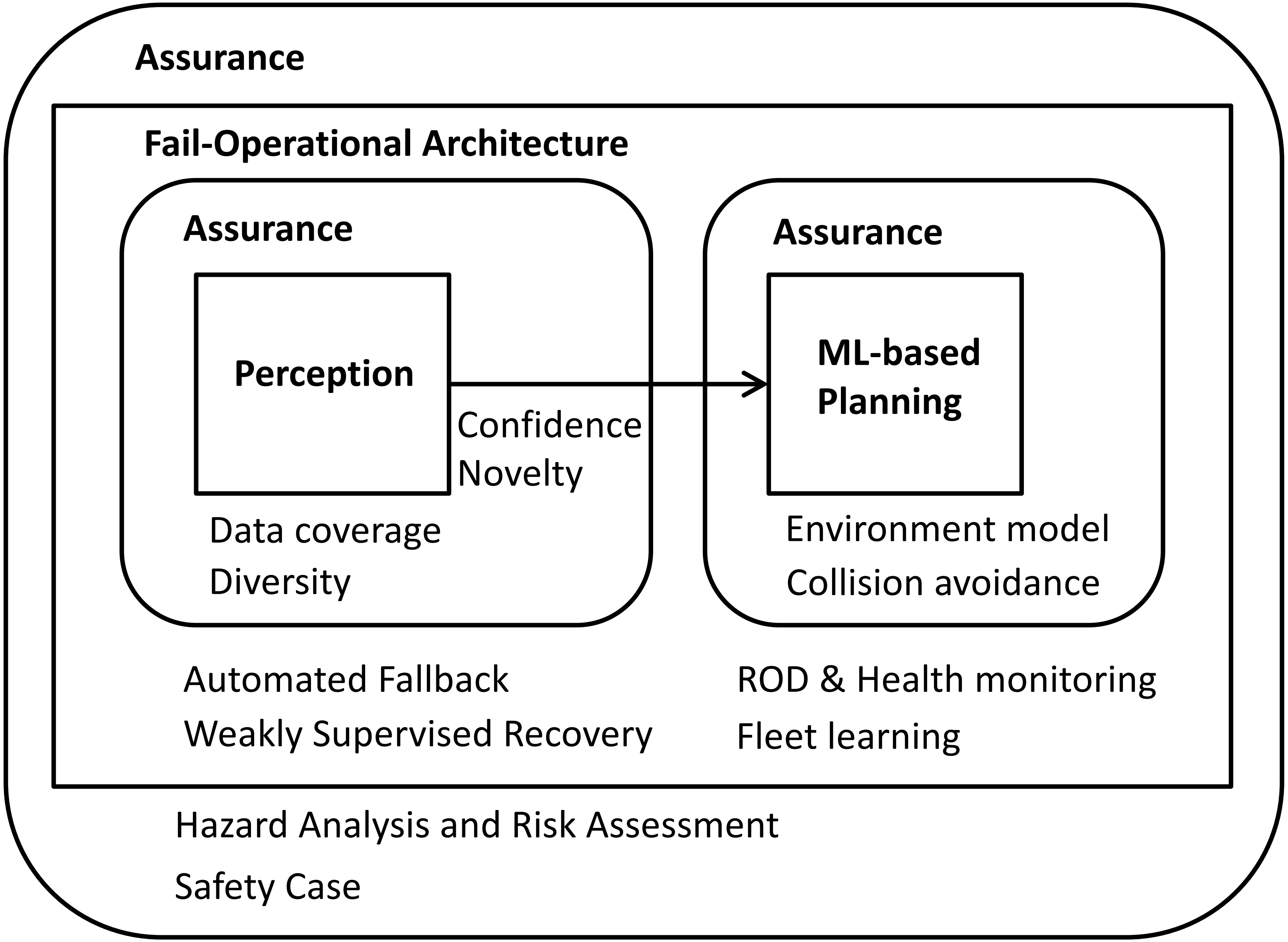
(ISO / PAS 21448)

- Sensor noise & limitations
- Machine learning errors
- Inadequate driving behavior
- DDT fallback failures

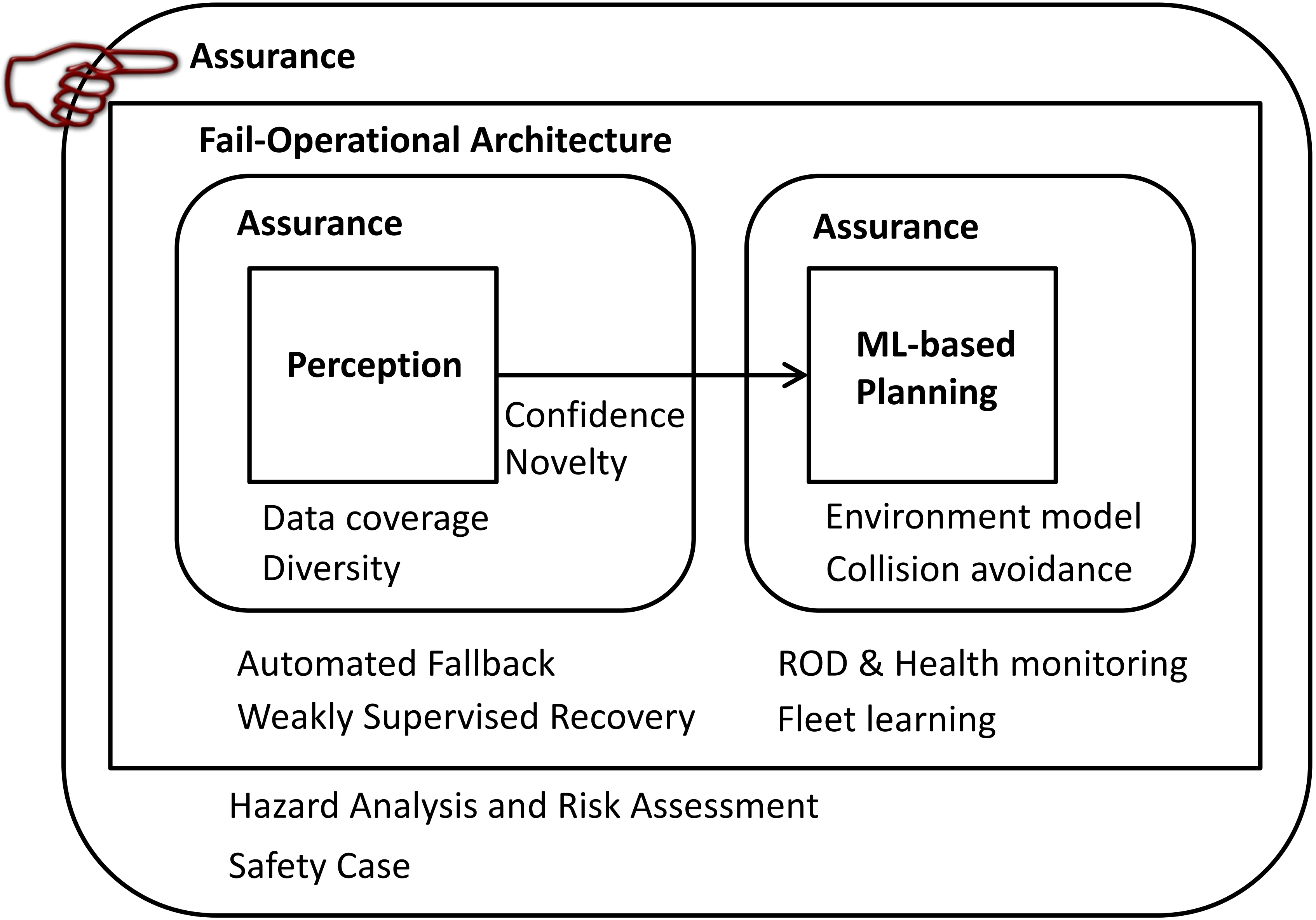
SAE J3061

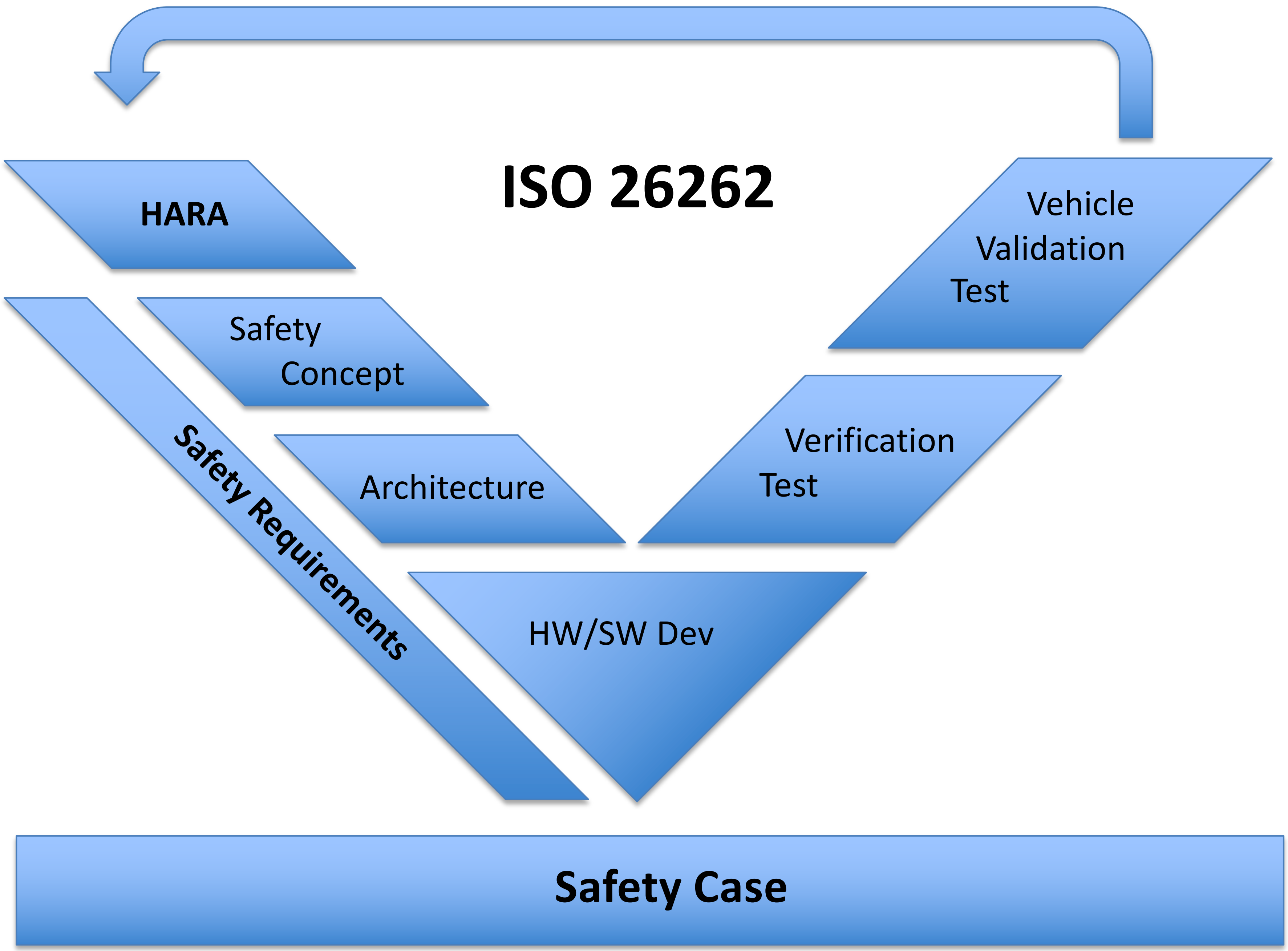
- Cyber attacks

LAVA: Learned & Assured Vehicle Autonomy

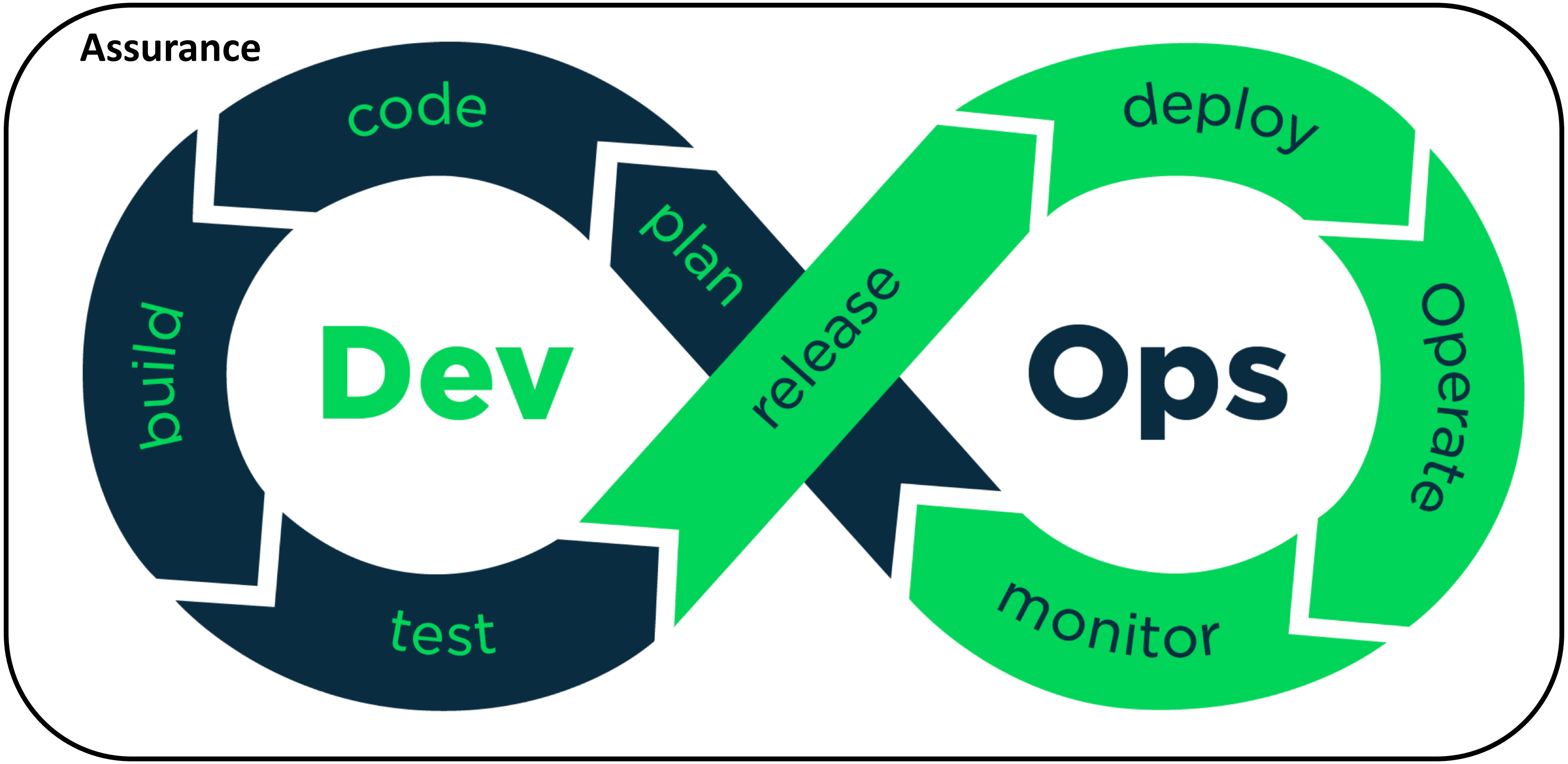


LAVA: Learned & Assured Vehicle Autonomy




HARA
ISO 26262
Vehicle Validation Test
Safety Concept
Safety Requirements
Architecture
Verification Test
HW/SW Dev
Safety Case

DevOps for ADS Software



Shadow testing

Design of experiments & fleet learning

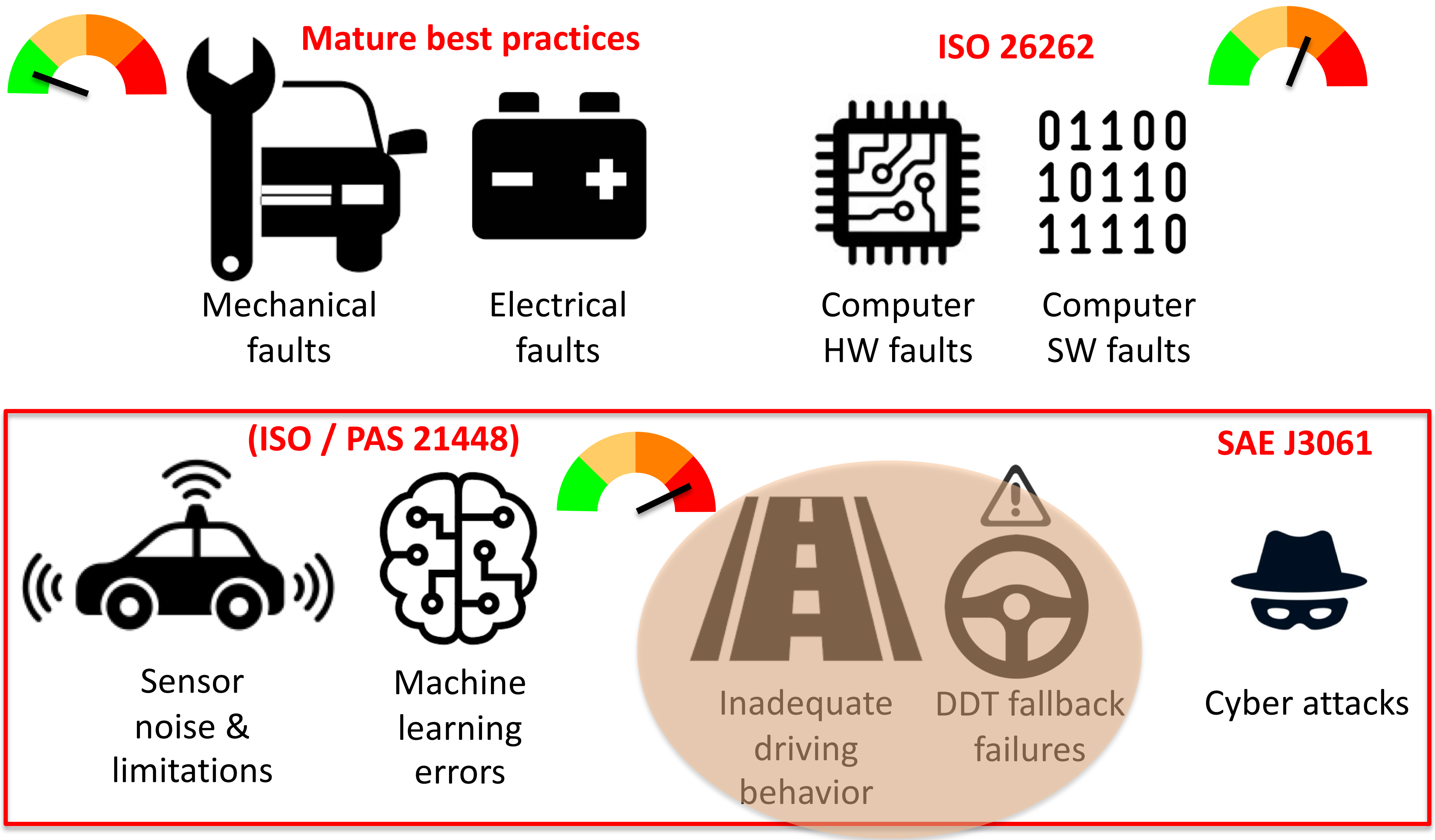
What field data to collect?

Update assurance

Incremental assurance

Safety case evolution

ADS Hazard Sources

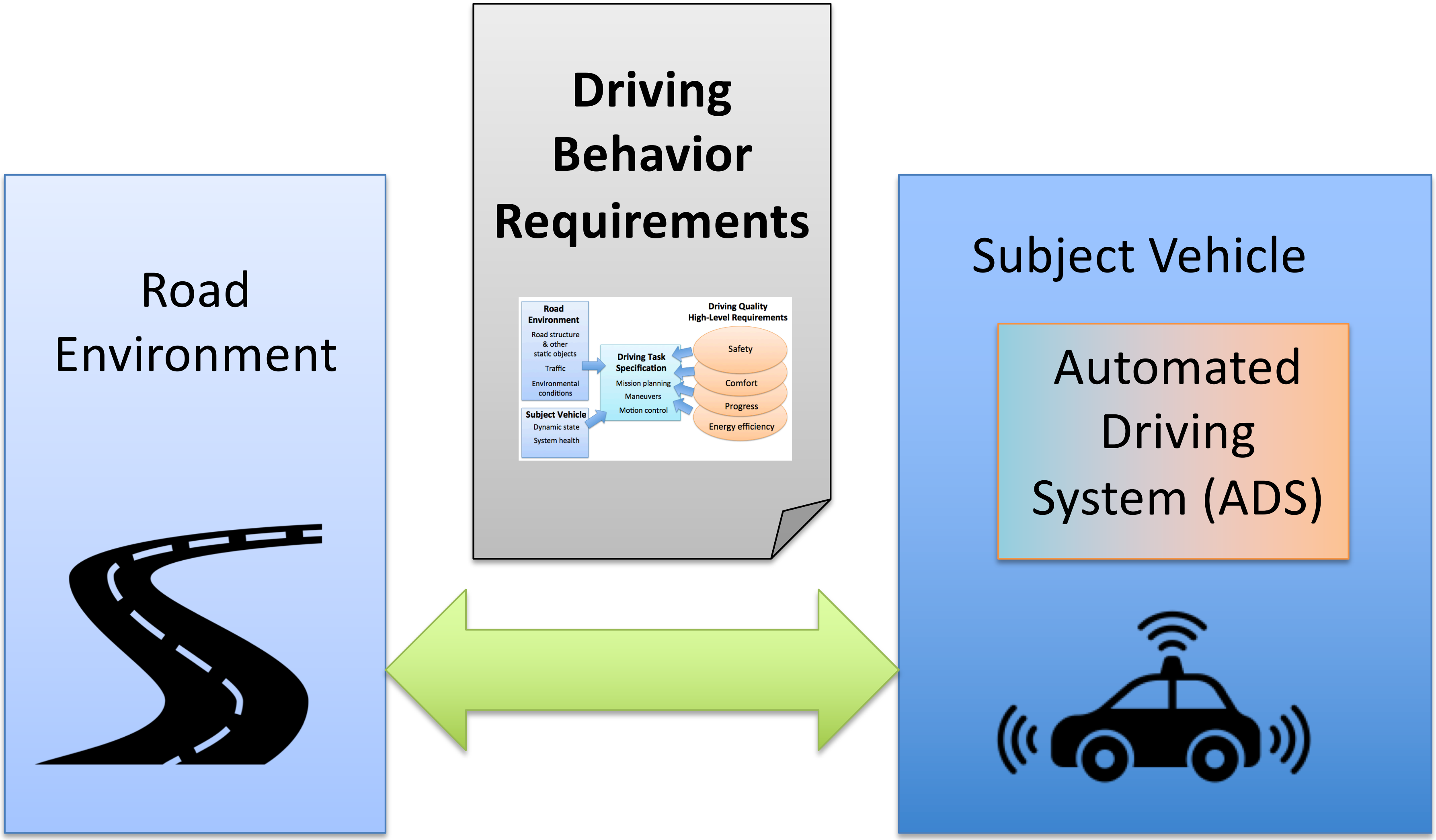


WISE Drive

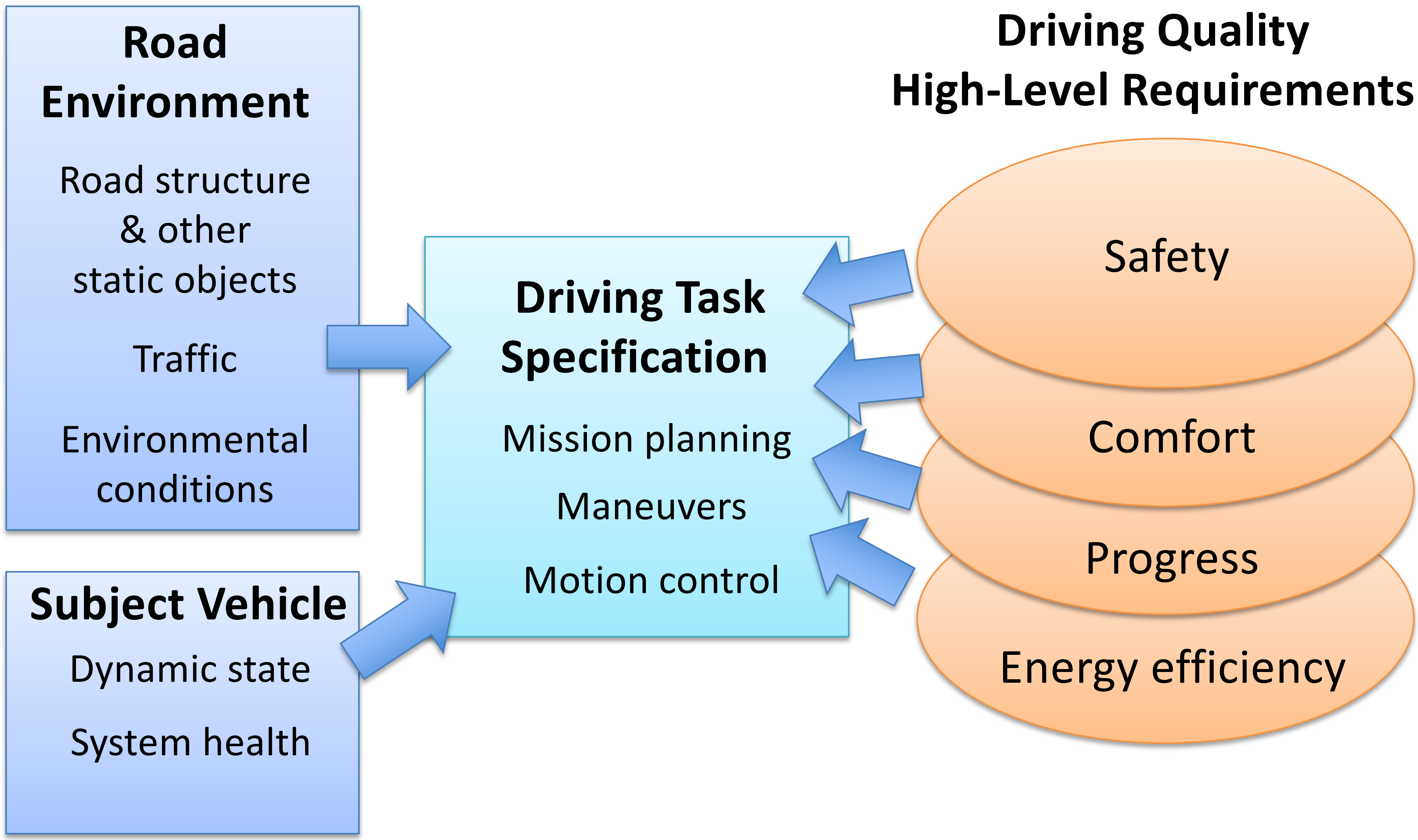
- Framework for analyzing and specifying requirements for an ADS
- Instantiated for a sample ODD on UW Moose
- Input into standardization (SAE J3164)

https://uwaterloo.ca/wise-lab/projects/wise-drive-requirements-analysis-framework-automated-driving

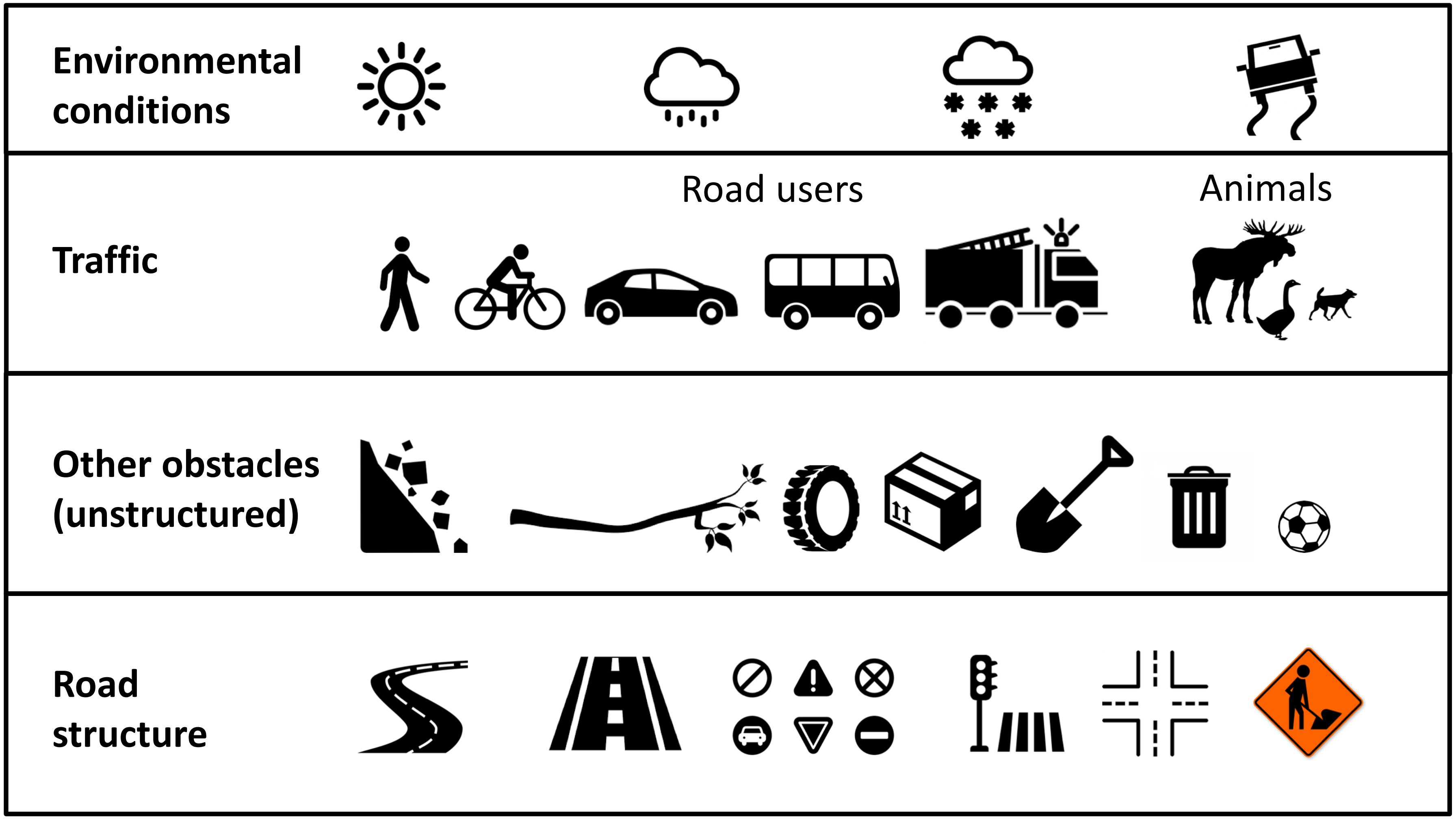
Requirements Specification



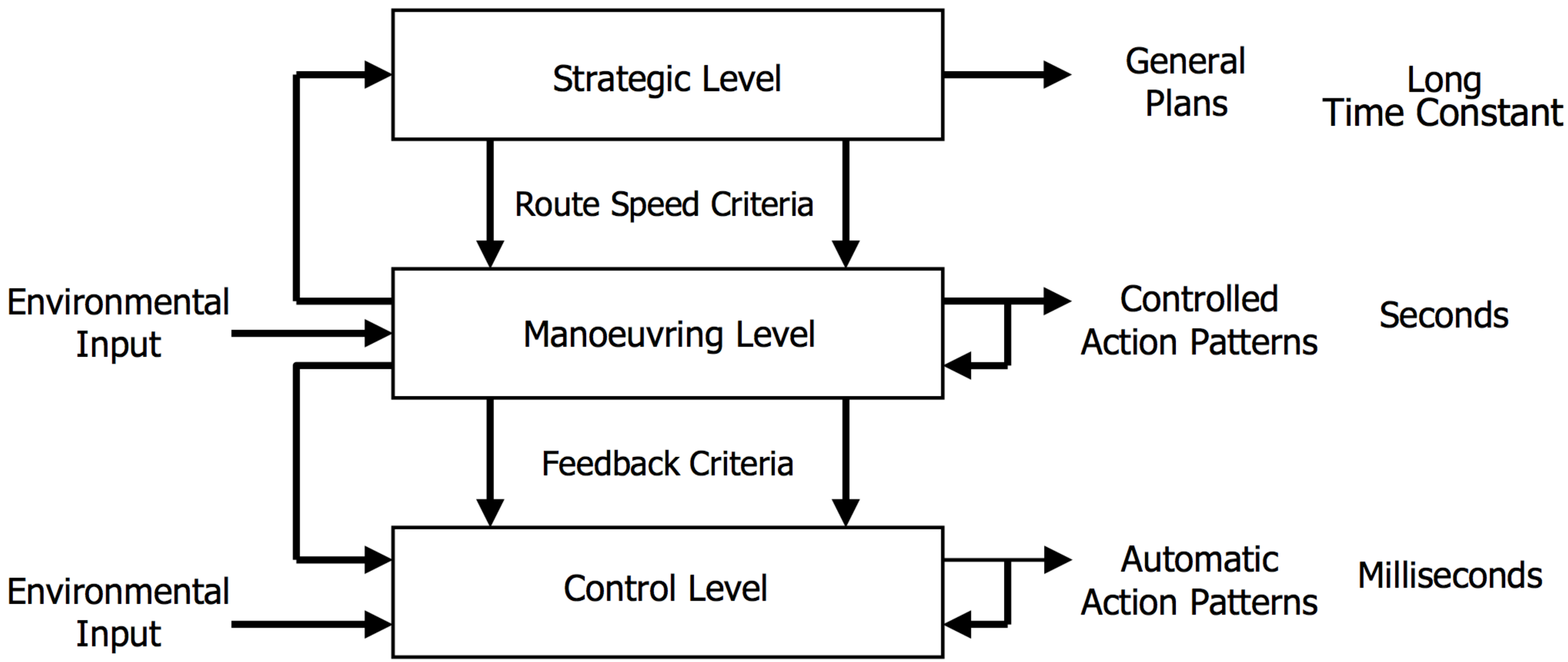
Driving Behavior Specification



Road Environment Ontology

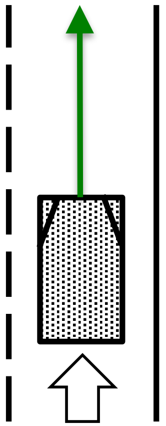


Driving Task

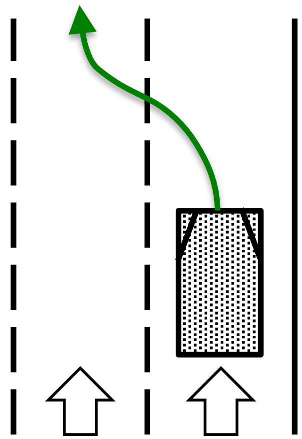


John Micheon, 1985

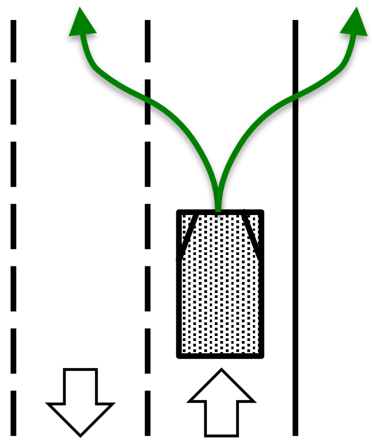
Primary Maneuvers



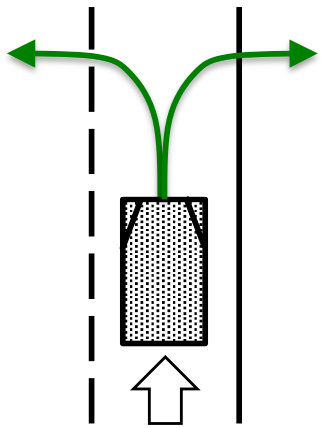
Lane maintenance



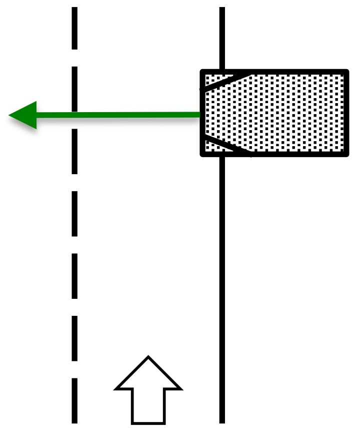
Lane change



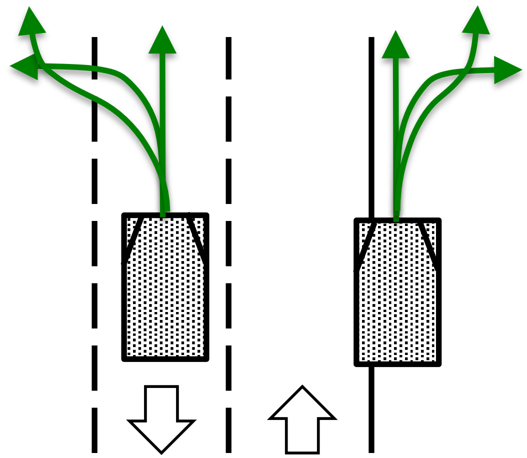
Swerve out of a same-direction traffic lane



Turn out of a same-direction traffic lane



Crossing a lane



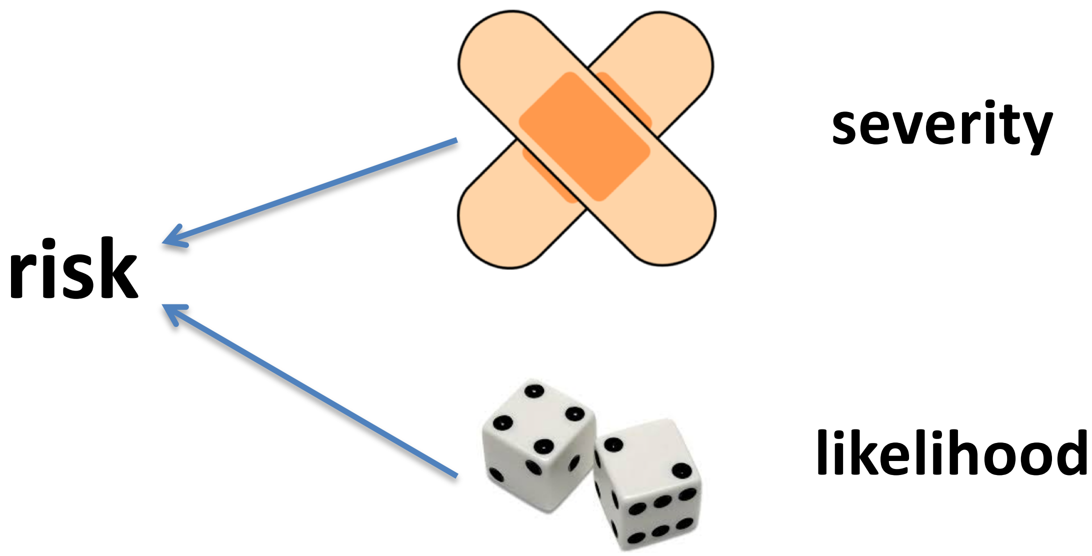
Movements outside a same-direction traffic lane

Secondary Maneuvers

1. Overtaking
2. Passing
3. Intersection handling
   – includes handling circular and non-circular intersections
4. Interchange handling
   – includes using acceleration lanes, entry and exit ramps, and weaving areas
5. Pedestrian crossing handling
6. Cycle crossing handling
7. Railway crossing handling
8. Turnabouts
9. Joining and leaving traffic

Safety

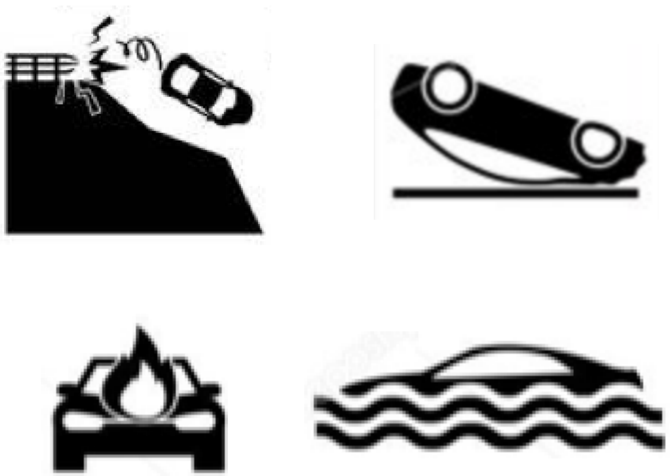
Absence of <u>unreasonable</u> risk of mishap



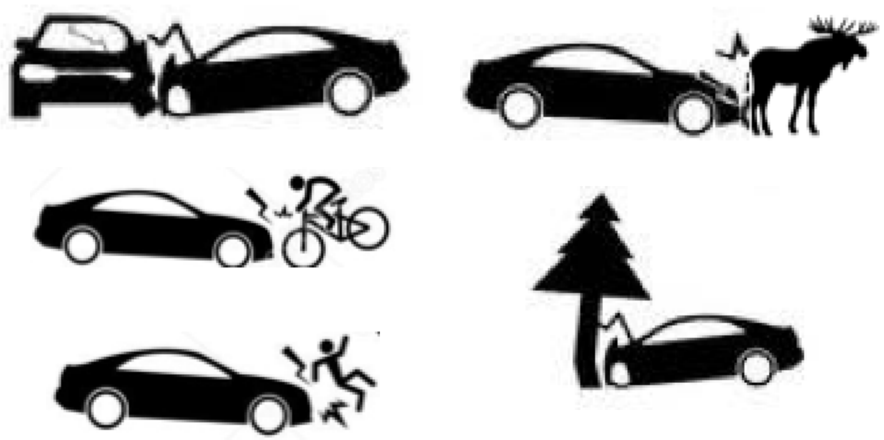
Driving Behavior Safety

Absence of unreasonable crash risk due to ADS driving behavior

Noncollisions



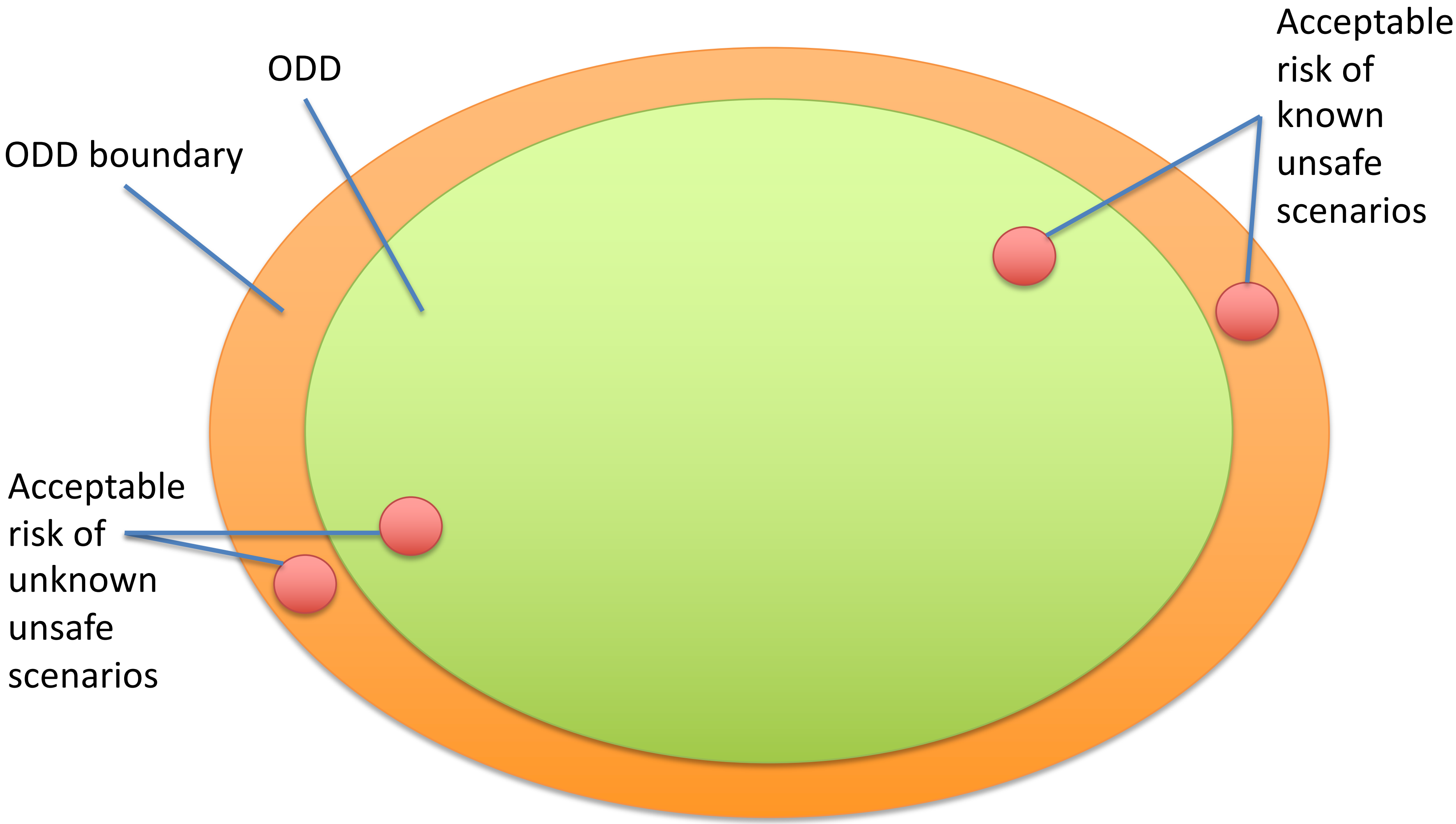
Collisions



Factors Influencing Risk Acceptability

- Risk level
- Risk reduction cost
- Benefit of the risky functionality (risk taking)
- Best practice (state of technology)
- Replacement risk
- Who controls risk
- Perception/public opinion

Assurance Target



Responsibility-Driven Safety

- Normal driving scenarios
  - Must not cause unacceptable risk increase
  - Low/high demand (incl. other road user errors)

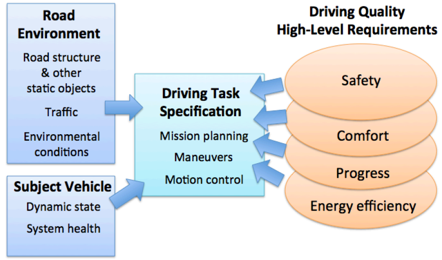
- Emergency scenarios
  - Near-crash
    - Must avoid crash if it can
  - Crash
    - Must mitigate if it can
    - Dilemmas often addressed by blame assignment
  - Fallback
    - Must minimize overall risk

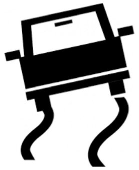
(related: Responsibility-Sensitive Safety, https://arxiv.org/pdf/1708.06374)

Blame vs. Injury Risk

GM Cruise Chevy vs. motorcycle crash
https://www.dmv.ca.gov/portal/wcm/connect/1877d019-d5f0-4c46-b472-78cfe289787d/GMCruise_120717.pdf?MOD=AJPERES

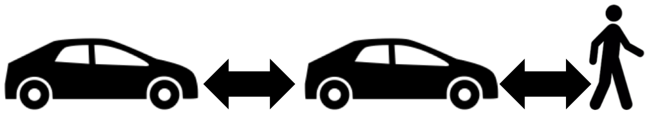
High-Level Behavior Safety Requirements (Normal Driving)





1. Vehicle stability

2. Assured clear distance ahead



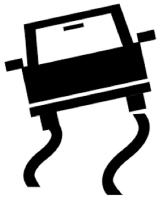
3. Minimum separation

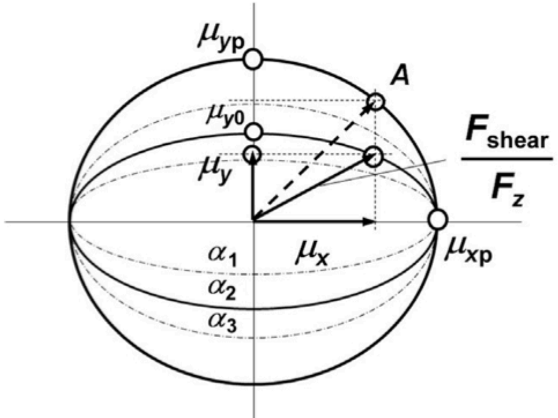
4. Traffic regulations

5. Informal traffic rules
   (best practices)

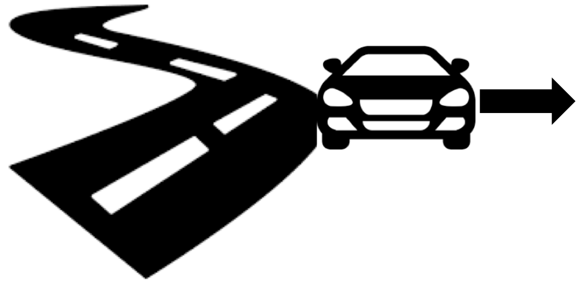
Behavioral Safety: 1. Vehicle Stability

Skid stability





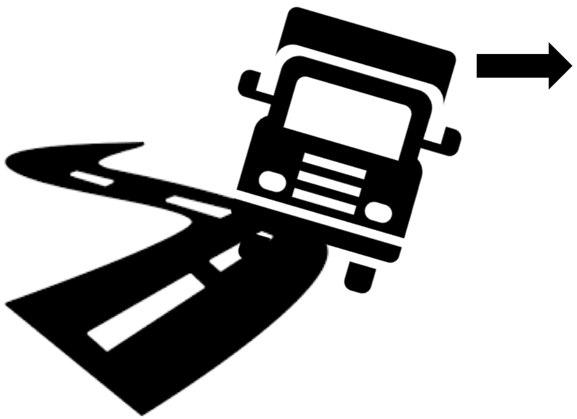
Friction ellipses



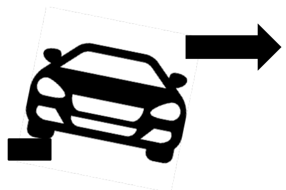
$e + \mu_y = v^2 / 127R$

Roll stability

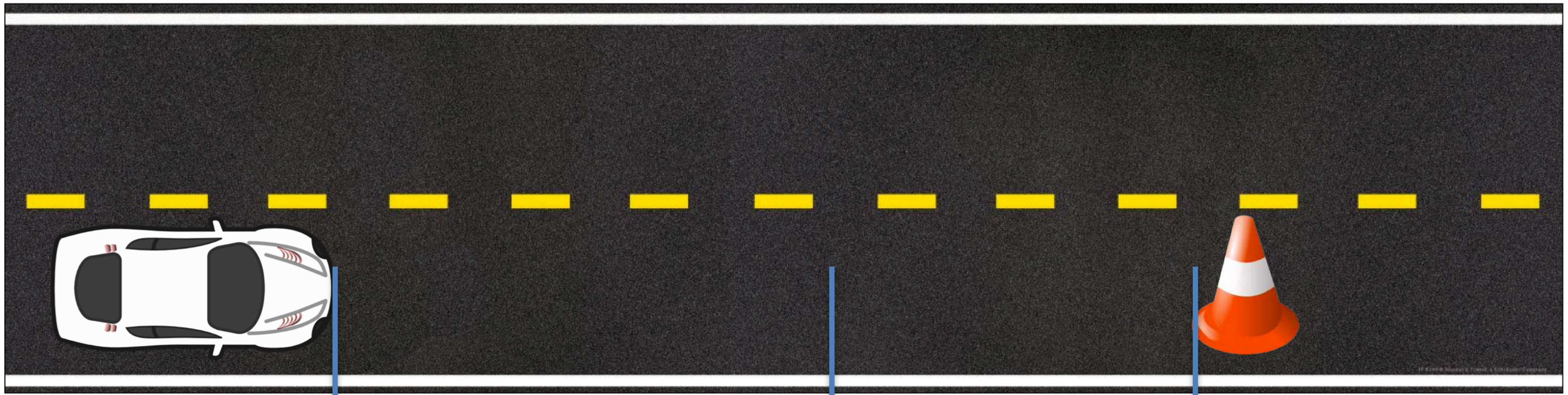
Untripped



Tripped



Behavioral Safety: 2. Assured Clear Distance Ahead (ACDA)



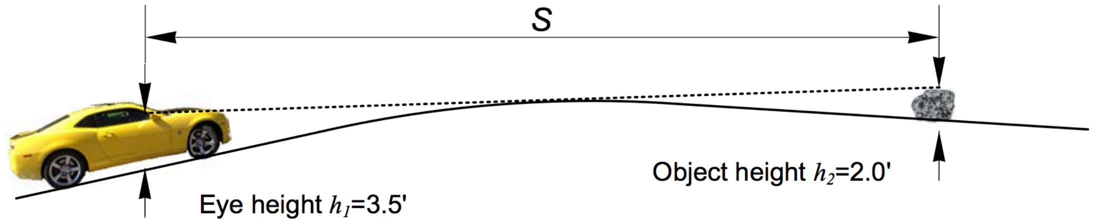
Stopping sight distance
(Perception-reaction time and braking distance)

Perception distance
(Range + road geometry)

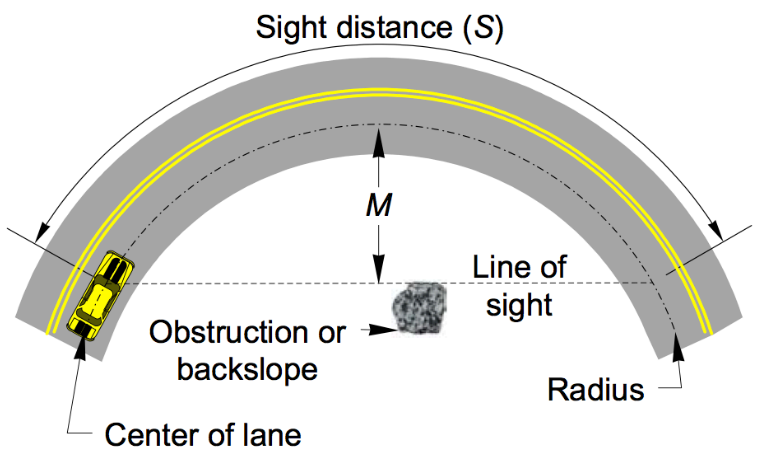
Limits safe speed

Behavioral Safety: 2. ACDA Perception Distance

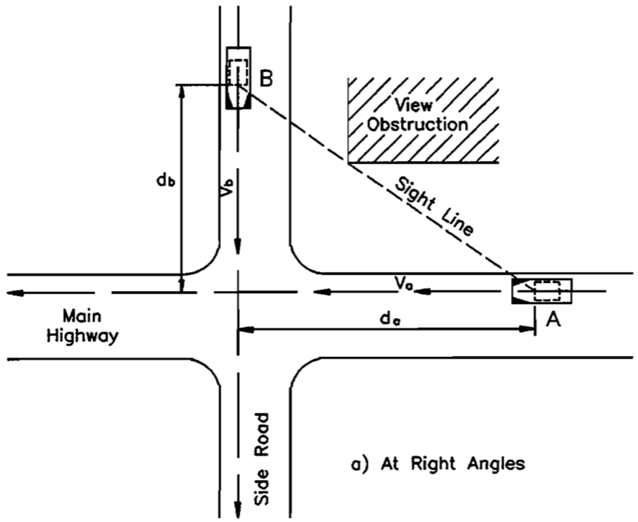
Crests



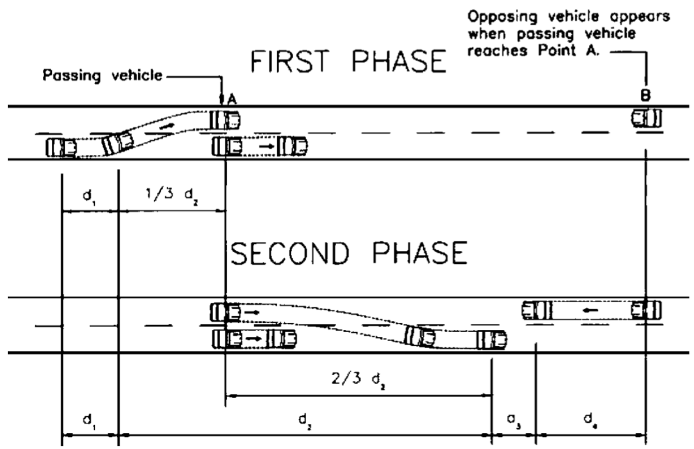
Curves



Intersections

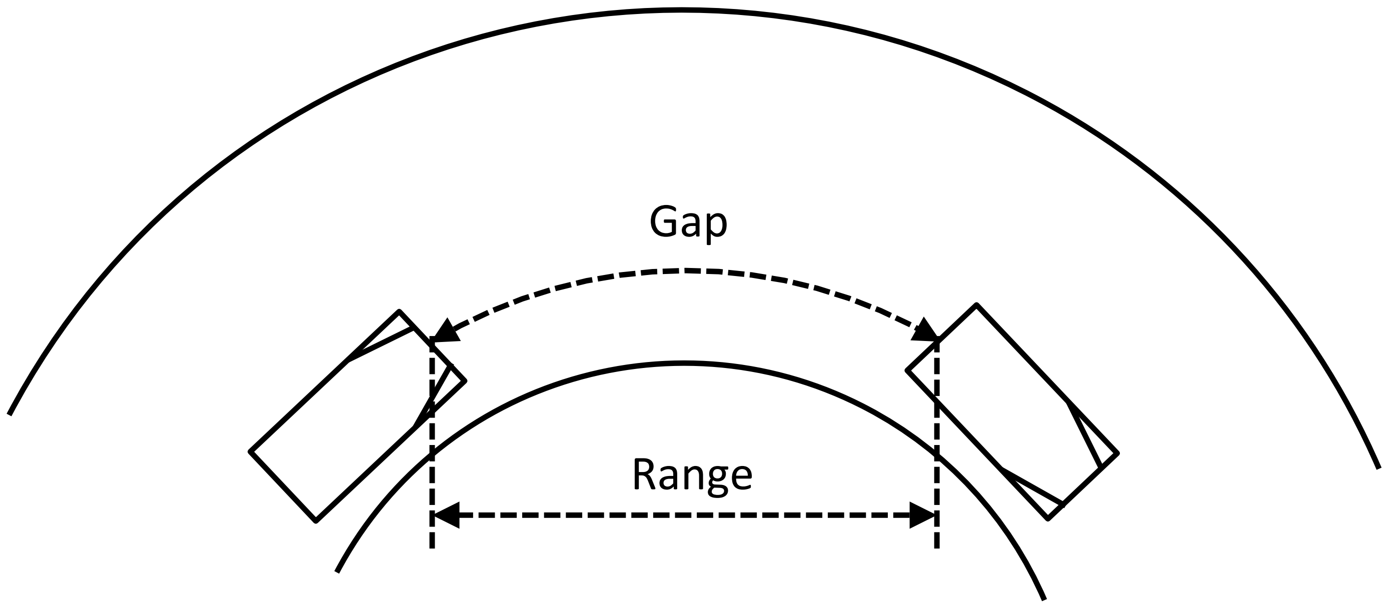


Overtaking



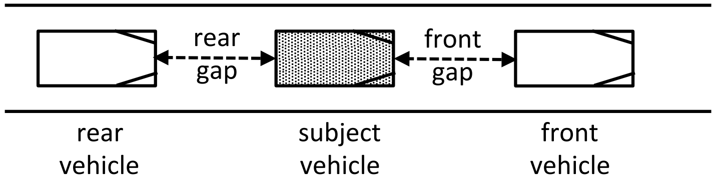
Behavioral Safety:
3. Minimum Separation

Separation in terms of **distance gap**, **time gap**, and **time-to-collision**

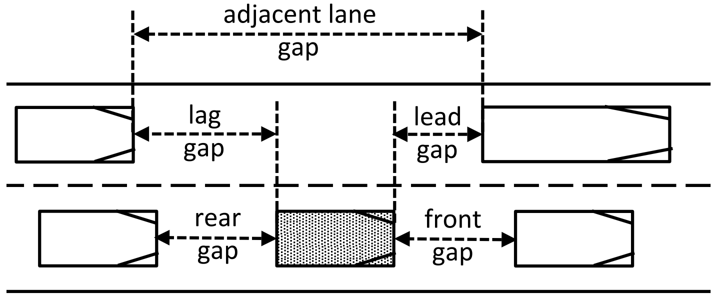


Behavioral Safety: 3. Minimum Separation Maneuver-Specific Gaps

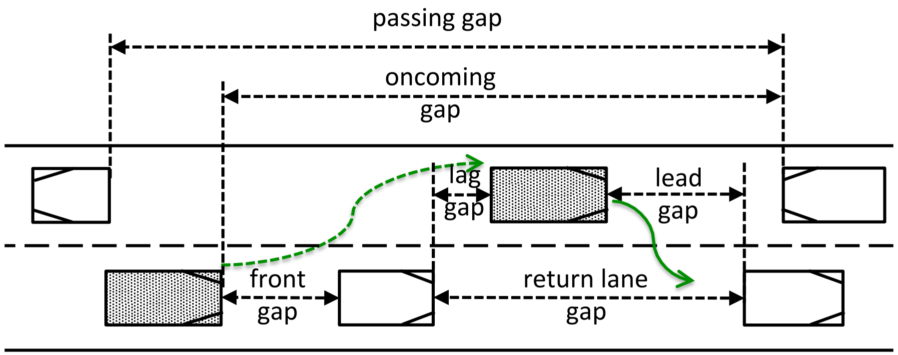
Car following



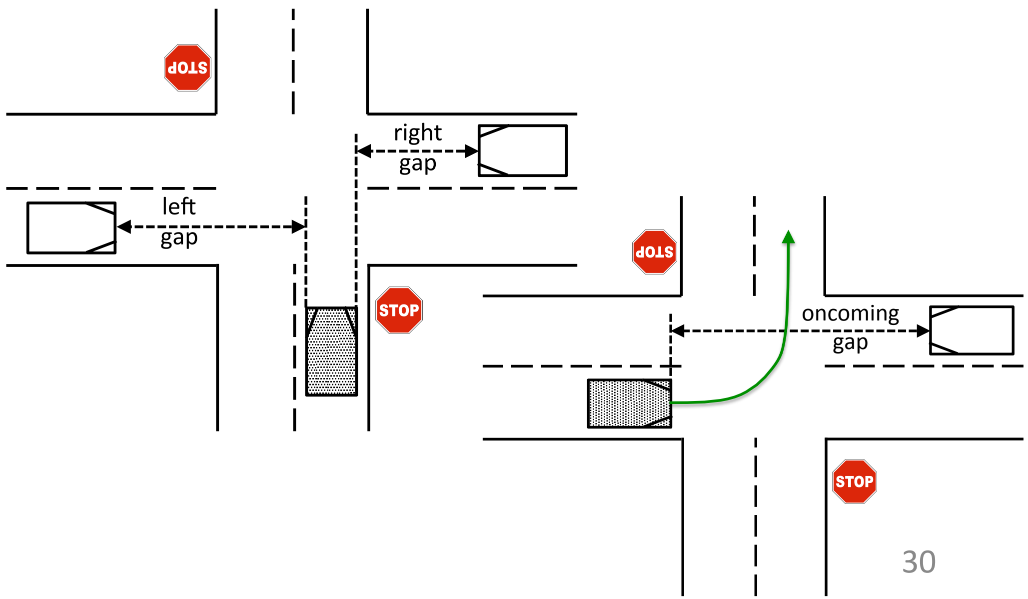
Lane changing



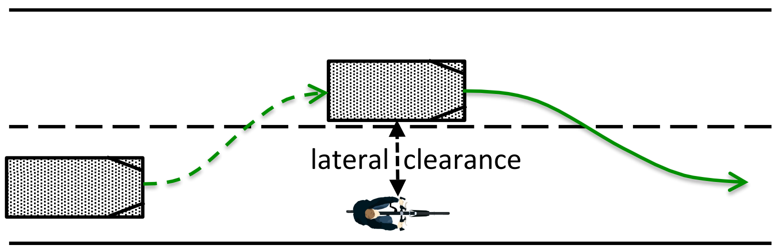
Overtaking



Intersection handling



Lateral clearance



Behavioral Safety:
4. Traffic Regulations

Safe speed (ACDA)

Yielding to other road users rules

Obeying regulatory traffic signs & signals

Where to drive

Reacting to emergency vehicles & school buses

U-turn prohibitions

Safe following gap

Passing rules

Signaling stops & turns

Parking restrictions

Use of passing beam

Required behavior at railway crossings

...

Behavioral Safety:
5. Informal Traffic Rules

2/3 – second rule

Responding to tailgaters

How early to signal turns

Delayed acceleration at signalized intersections

Lane selection

Anticipating aberrant behaviors of other road users

Responding to animals on the roadway

…

WISE Drive Documentation

WISE Drive comes with comprehensive documentation (over 350 pages) available from this page.

All eight documents in two zip archives: zip1, zip2

Driving Task Specification

Maneuver Catalog

K. Czarnecki. Automated Driving System (ADS) Task Analysis – Part 2: Structured Road Maneuvers. Waterloo Intelligent Systems Engineering Lab (WISE) Report, University of Waterloo, 2018, DOI: 10.13140/RG.2.2.23280.76800

Basic Motion Control Task Catalog

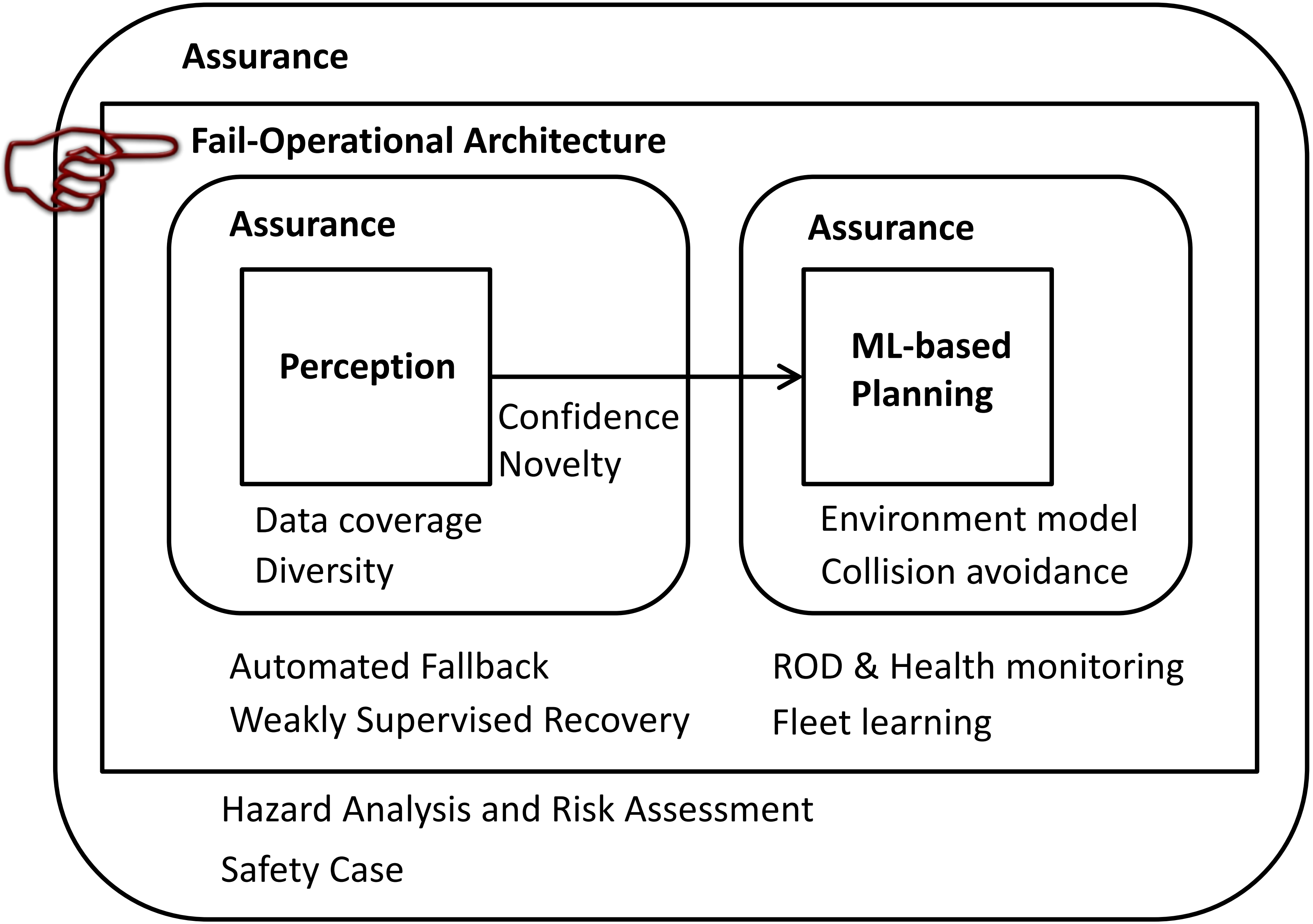
K. Czarnecki. Automated Driving System (ADS) Task Analysis – Part 1: Basic Motion Control Tasks. Waterloo Intelligent Systems Engineering Lab (WISE) Report, University of Waterloo, 2018, DOI: 10.13140/RG.2.2.29991.65447

Road Environment Specification

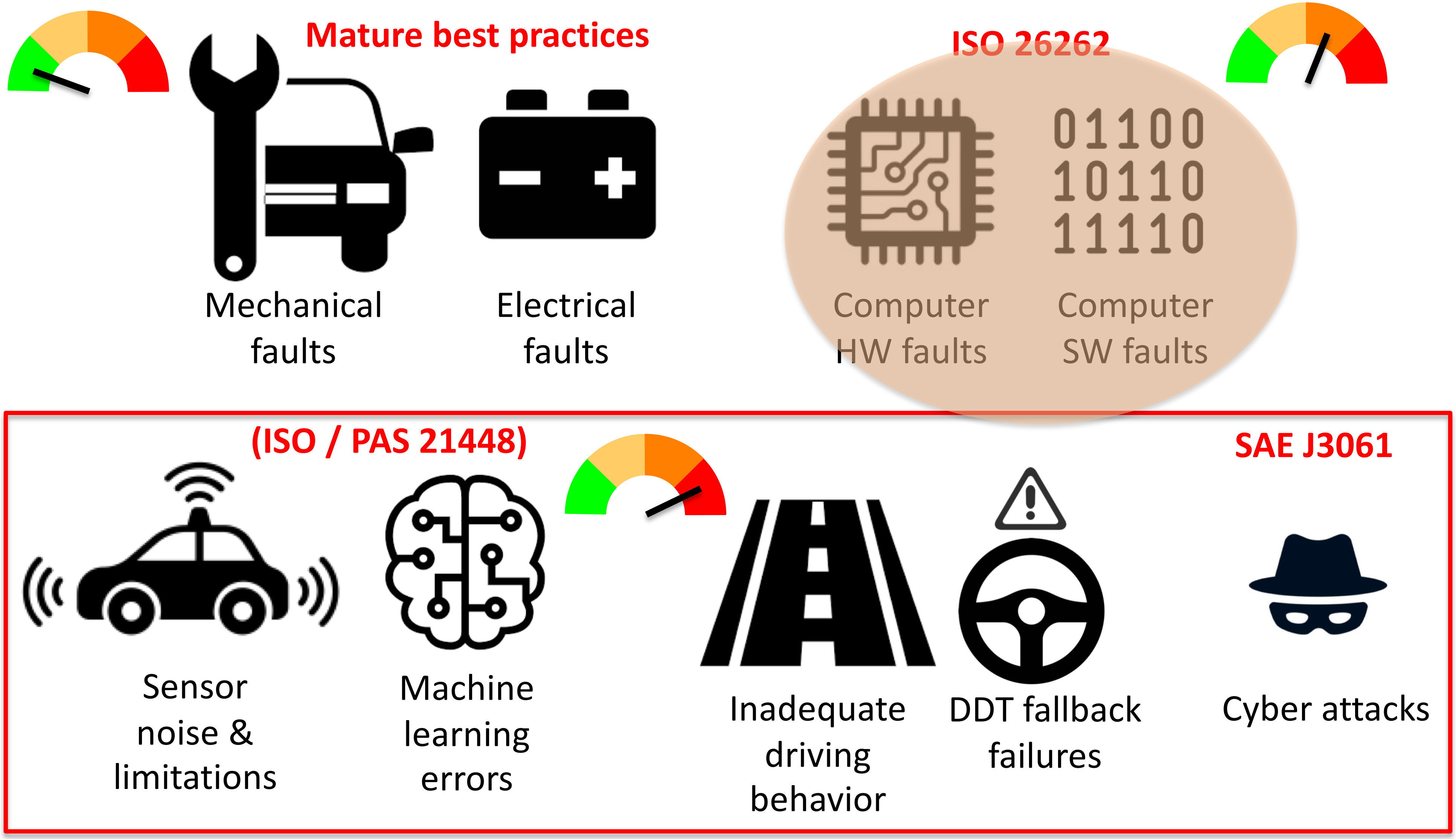
ODD Taxonomy

K. Czarnecki. Operational Design Domain for Automated Driving Systems – Taxonomy of Basic Terms. Waterloo Intelligent Systems Engineering Lab (WISE) Report, University of Waterloo, 2018, DOI:

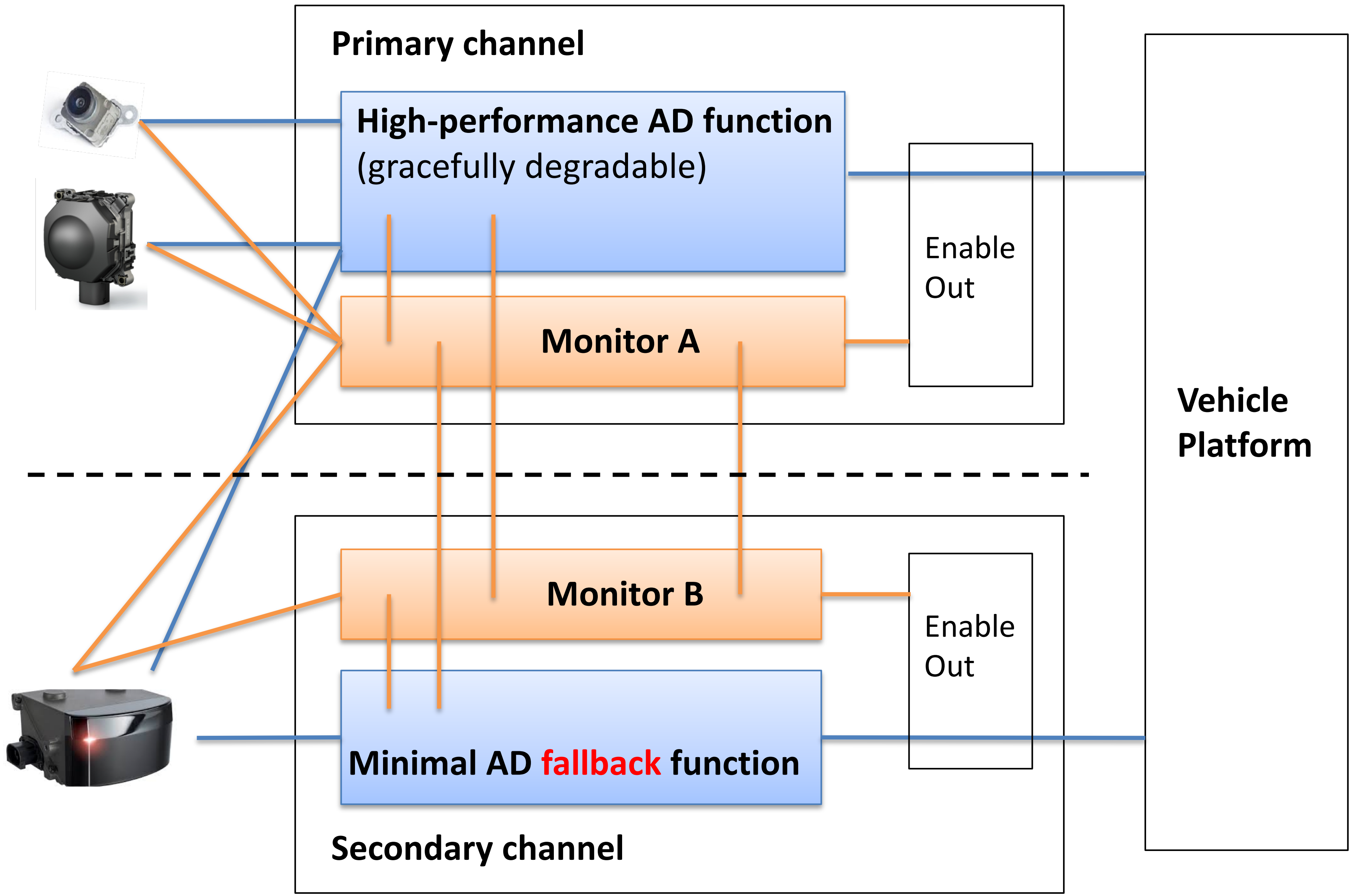
LAVA: Learned & Assured Vehicle Autonomy



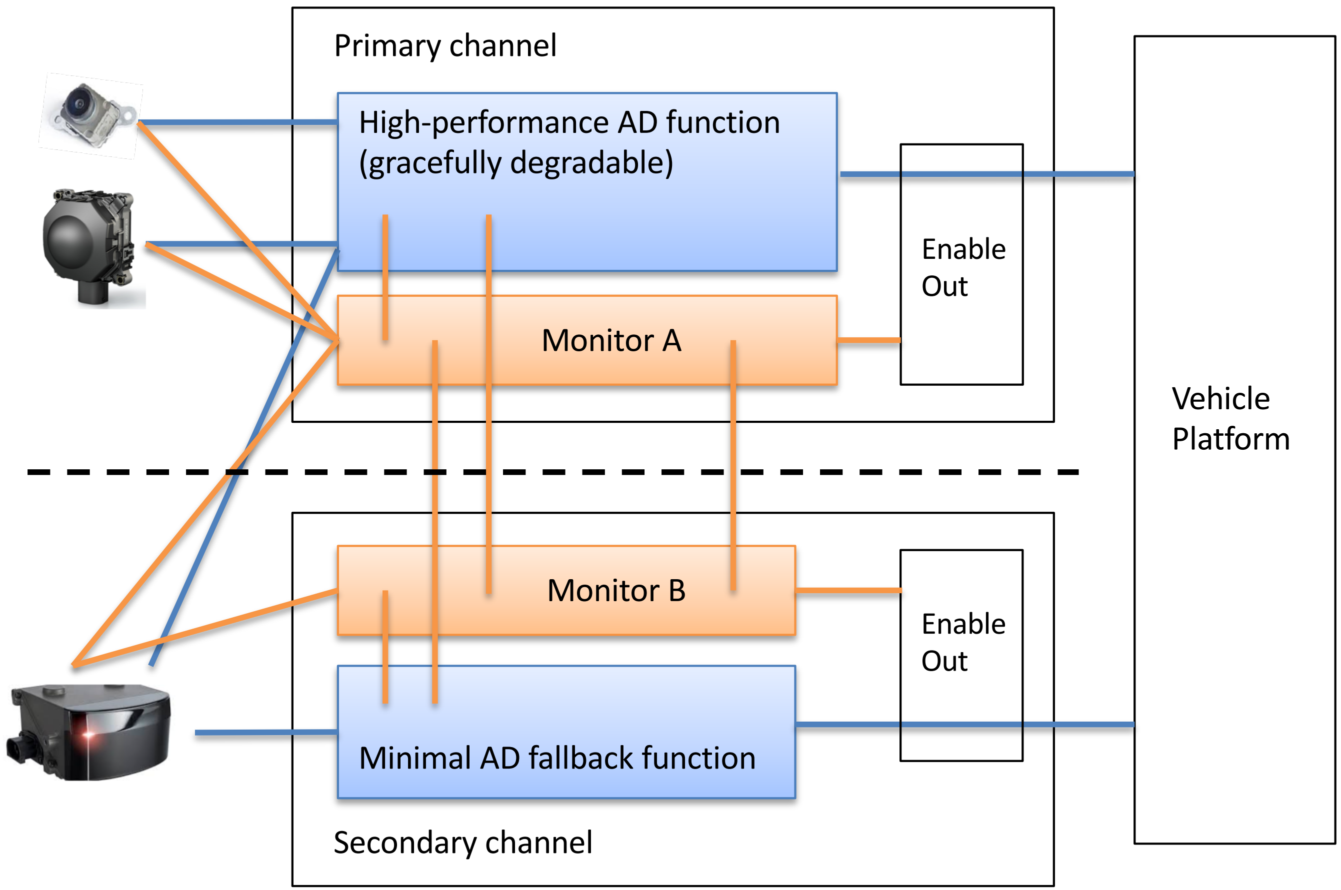
ADS Hazard Sources



Fail-Operational ADS Architecture



Fail-Operational ADS Architecture

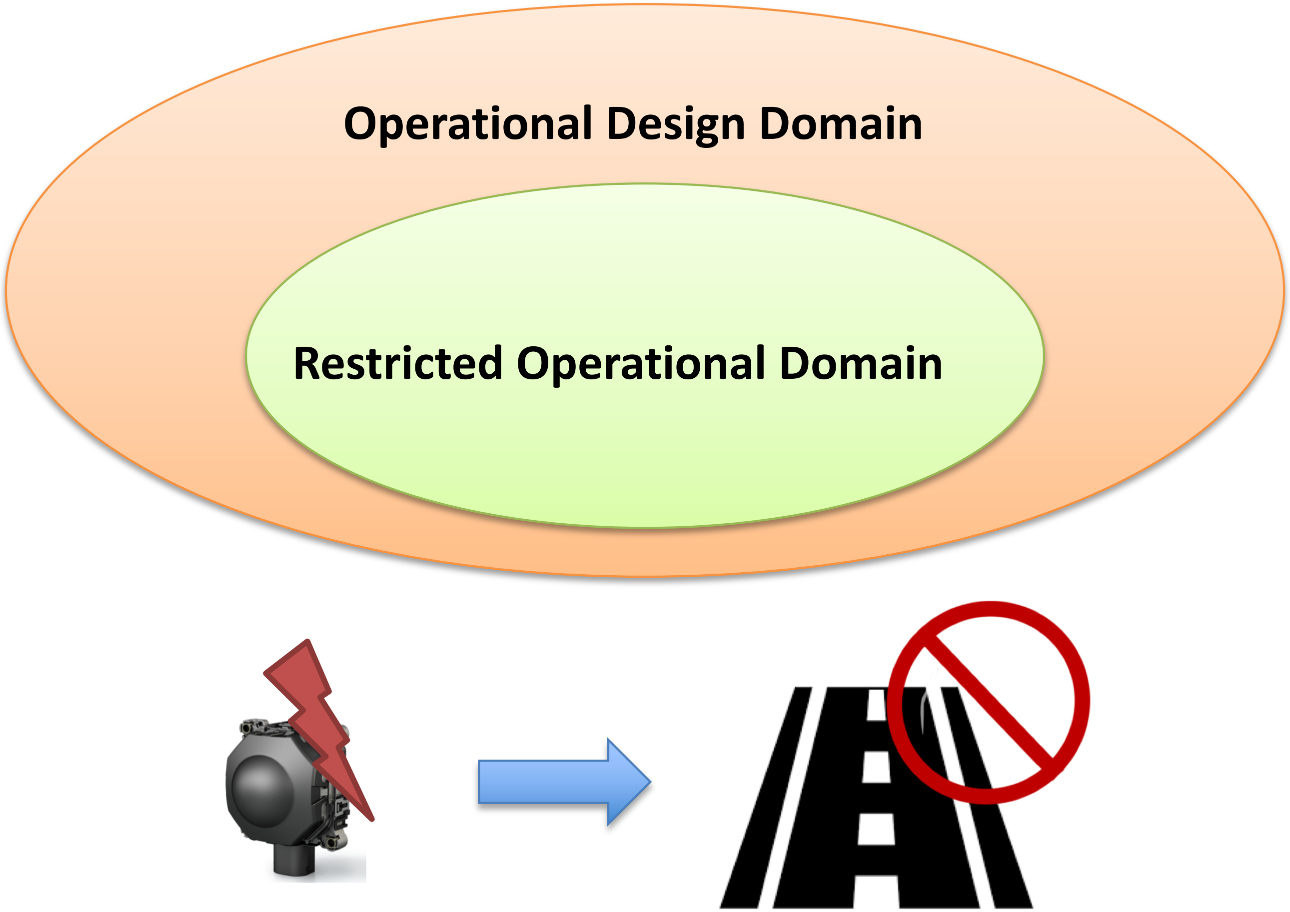


Dependability patterns:

- Redundancy
- Diversity
- Simplex
- Graceful degradation
- Monitoring of monitoring
- Minimized cost

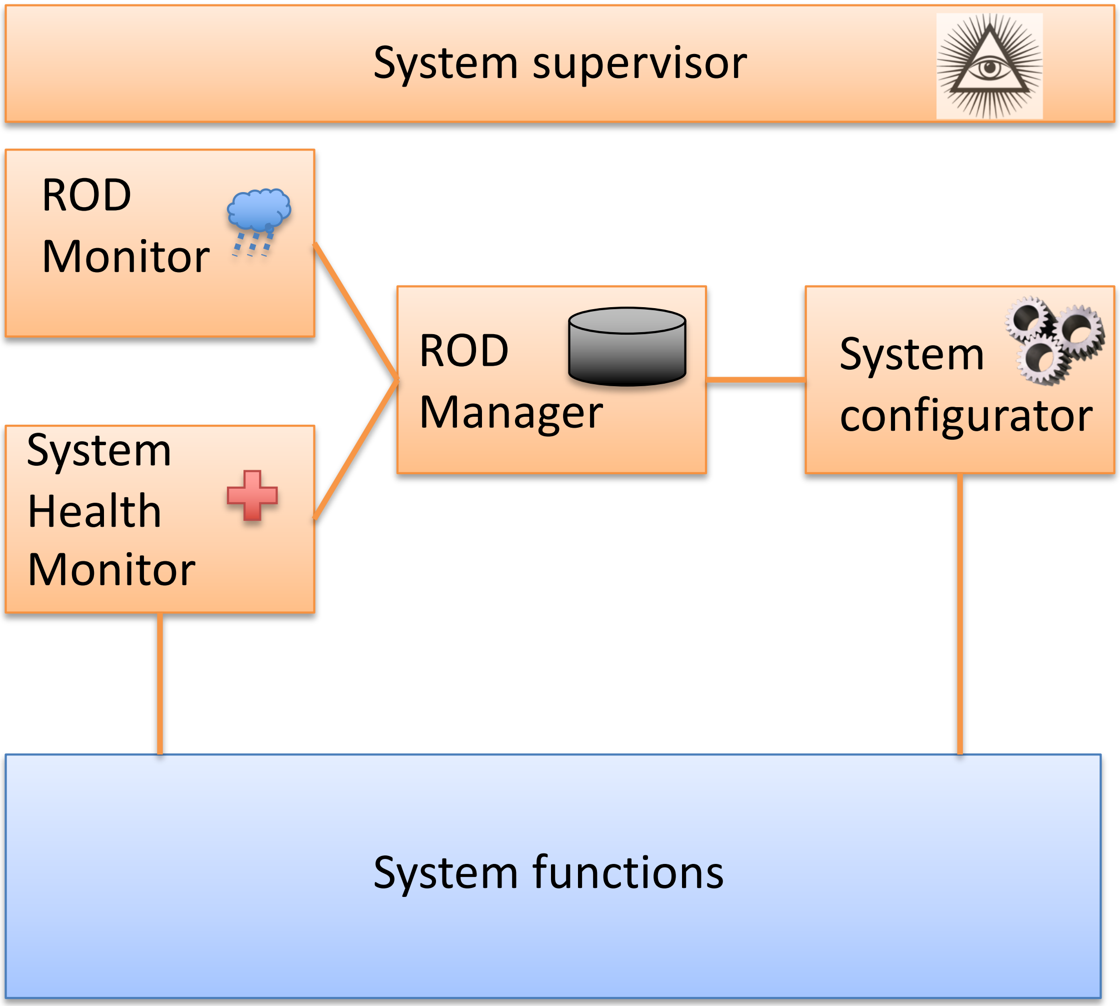
No single-point failures

ODD vs. ROD

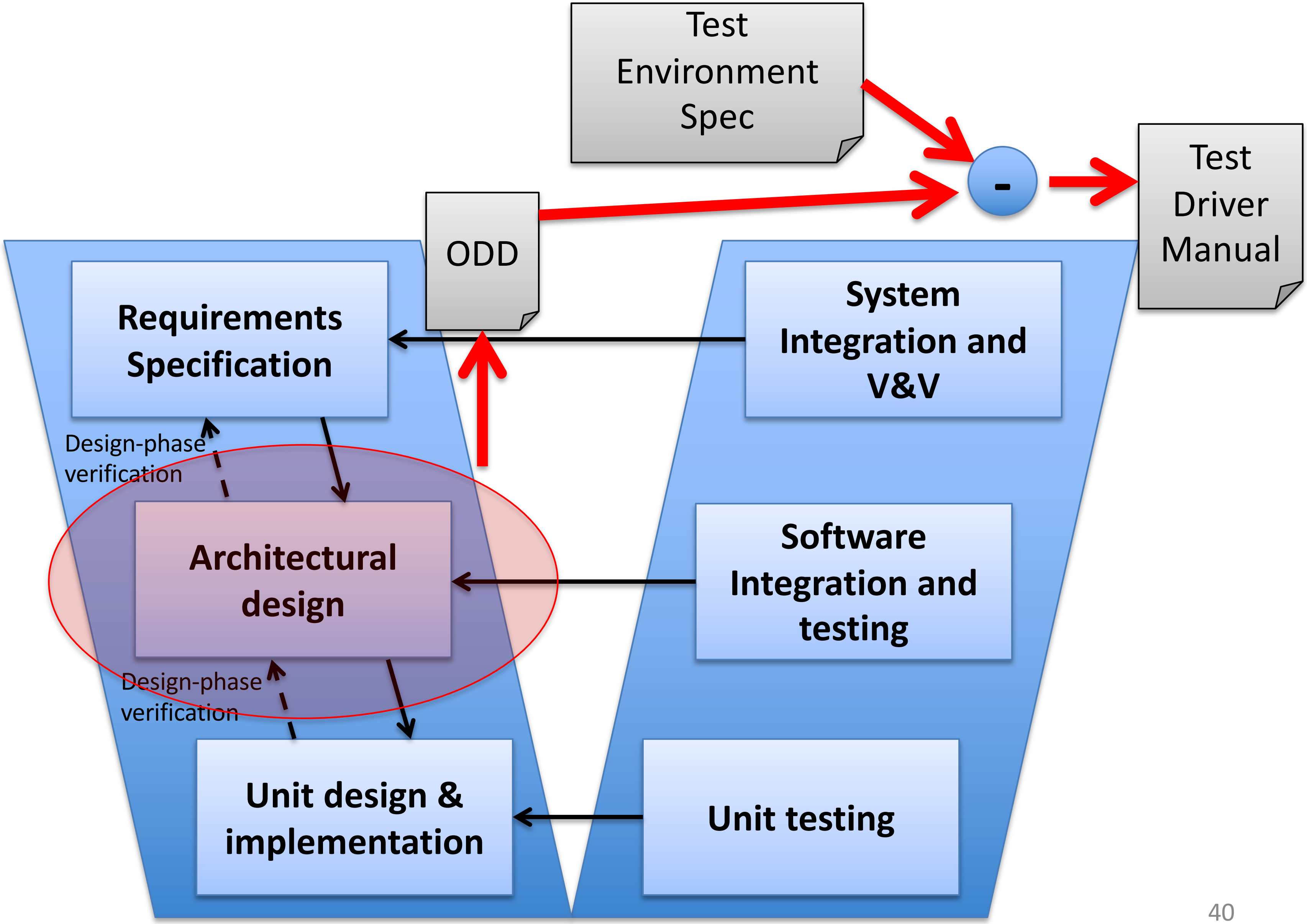


I Colwell, B Phan, S Saleem, R Salay, K Czarnecki. An Automated Vehicle Safety Concept Based on Runtime Restriction of the Operational Design Domain. IEEE Intelligent Vehicles Symposium (IV), 2018

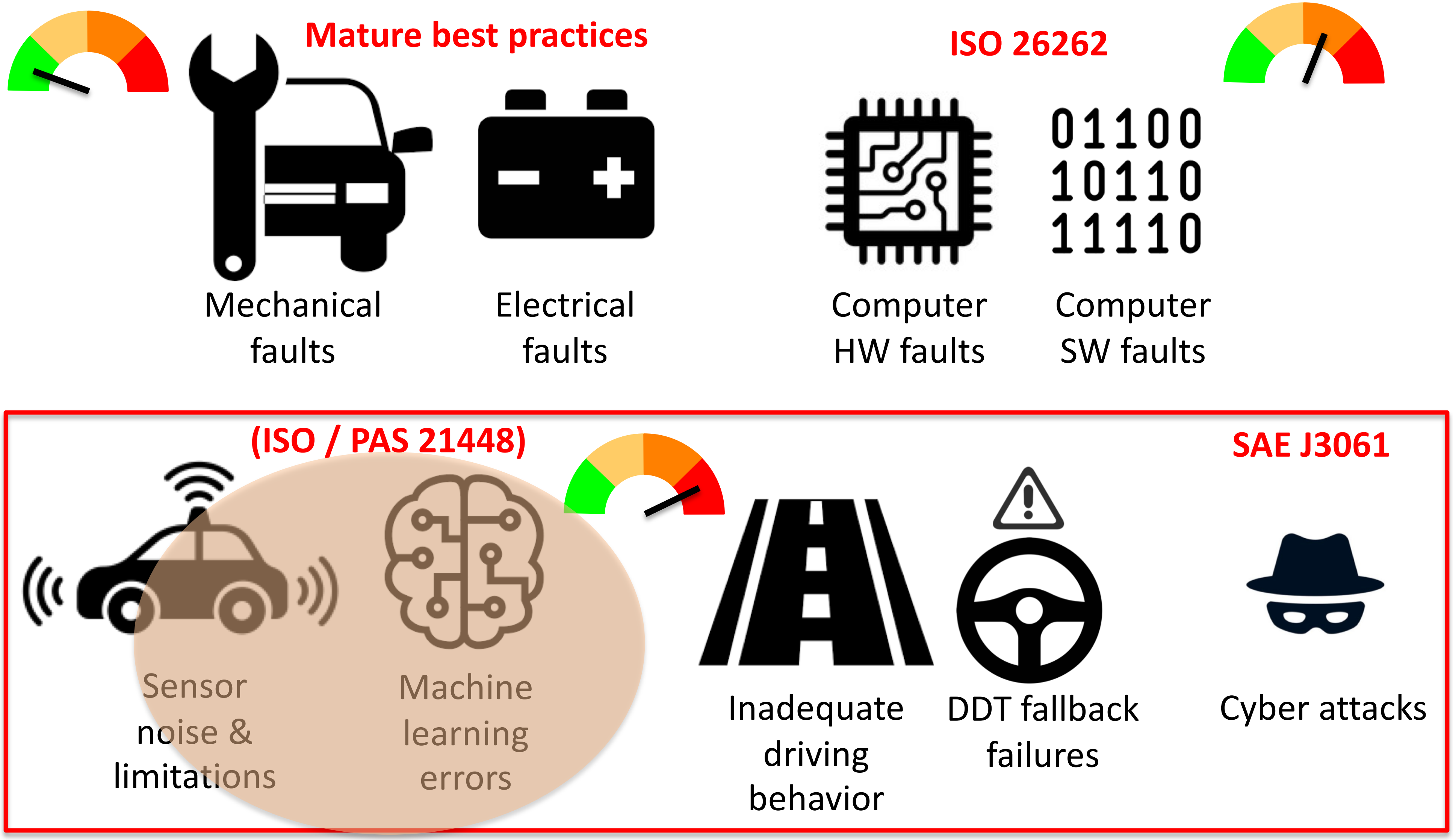
ROD Monitoring for Graceful Degradation



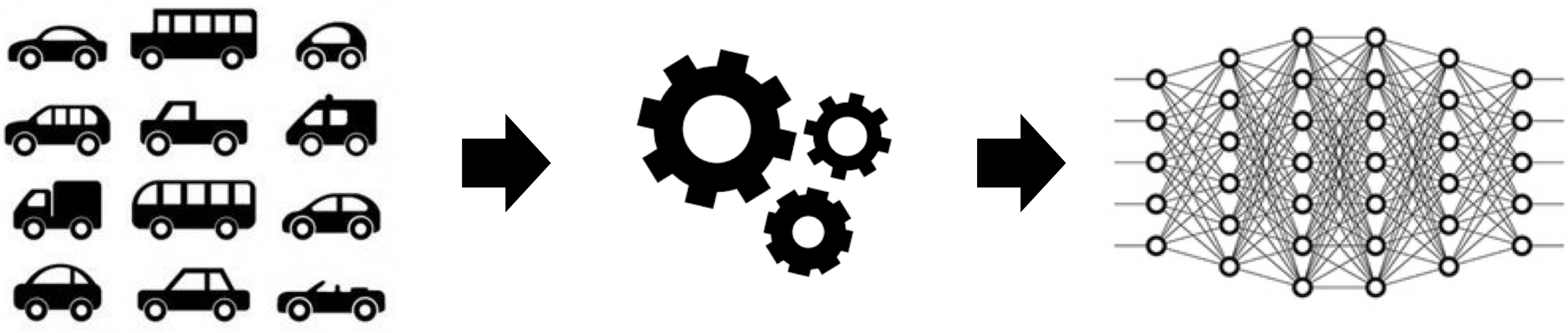
ODD vs. Test Environment



ADS Hazard Sources



Challenges of Assuring Machine Learned Components



Lack of specification

Lack of inspectability

R. Salay, R. Queiroz, K. Czarnecki. An Analysis of ISO 26262: Machine Learning and Safety in Automotive Software. SAE, 2018-01-1075, 2018; preliminary version also available at https://arxiv.org/abs/1709.02435

Lack of Complete Spec Affects Verification and Testing

Best practices

Spec notations

Design guidelines

Coding guidelines

Fault tolerance

Error detection & handling

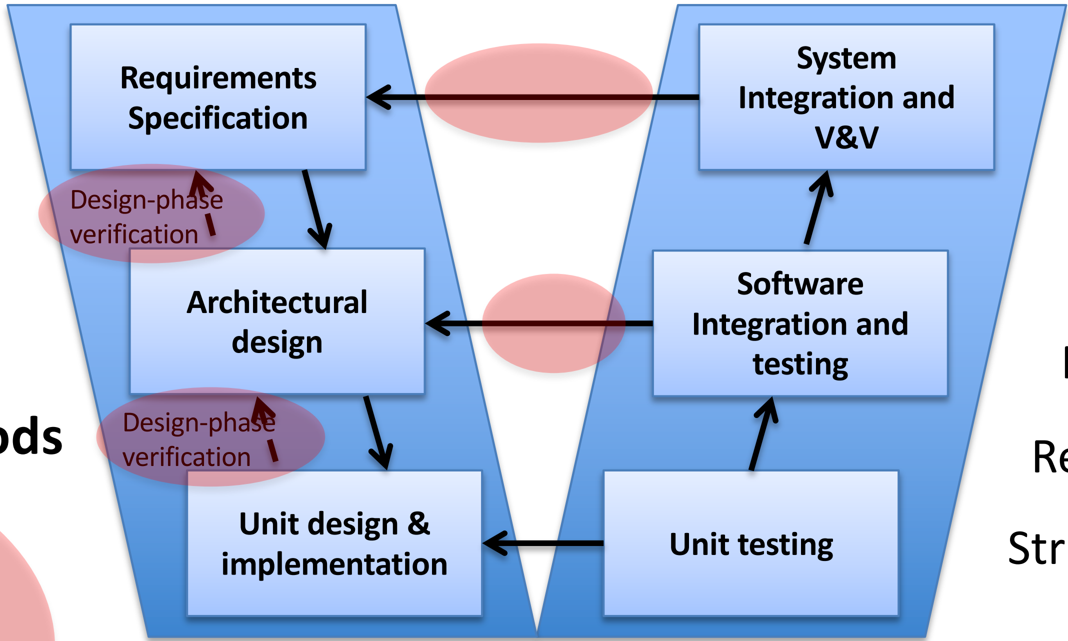
Verification methods

Walkthroughs

Inspections

Formal verification

Static code analysis



ISO 26262 Part 6

Testing methods

Requirements-based testing

Error guessing

Interface test

Fault injection test

Resource usage test

Structural coverage

Key Recommendations

- Partial specifications
  - Assumptions, necessary/sufficient conditions, in- and eqivariants
  - Runtime monitoring, test generation, regularization
- Data requirements
  - Domain coverage (e.g., ontology)
  - Risk profiling

56 pages

Using Machine Learning Safely in Automotive Software:

An Assessment and Adaption of Software Process Requirements in ISO 26262

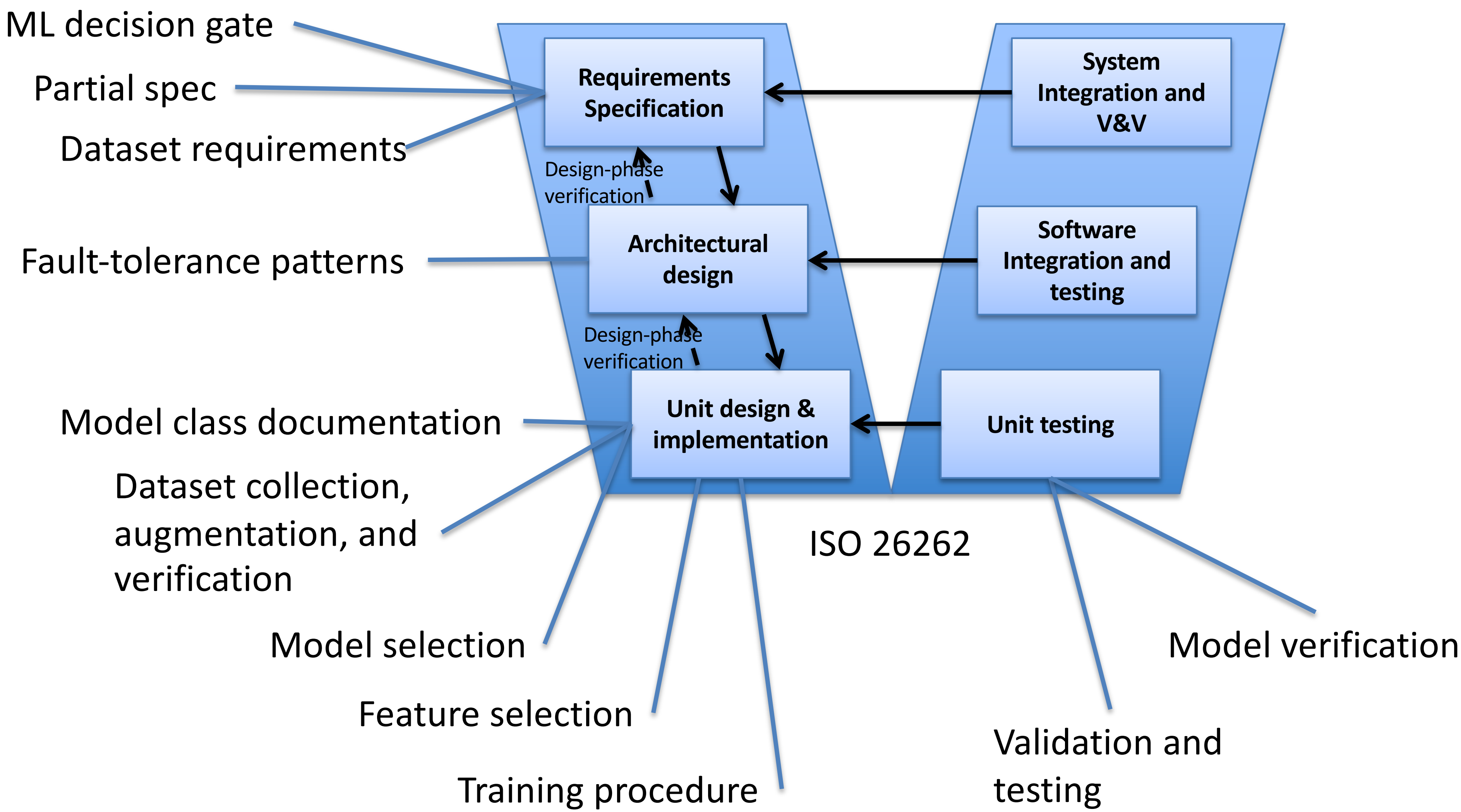
Rick Salay and Krzysztof Czarnecki
Waterloo Intelligent Systems Engineering (WISE) Lab
University of Waterloo
Canada

August 3, 2018

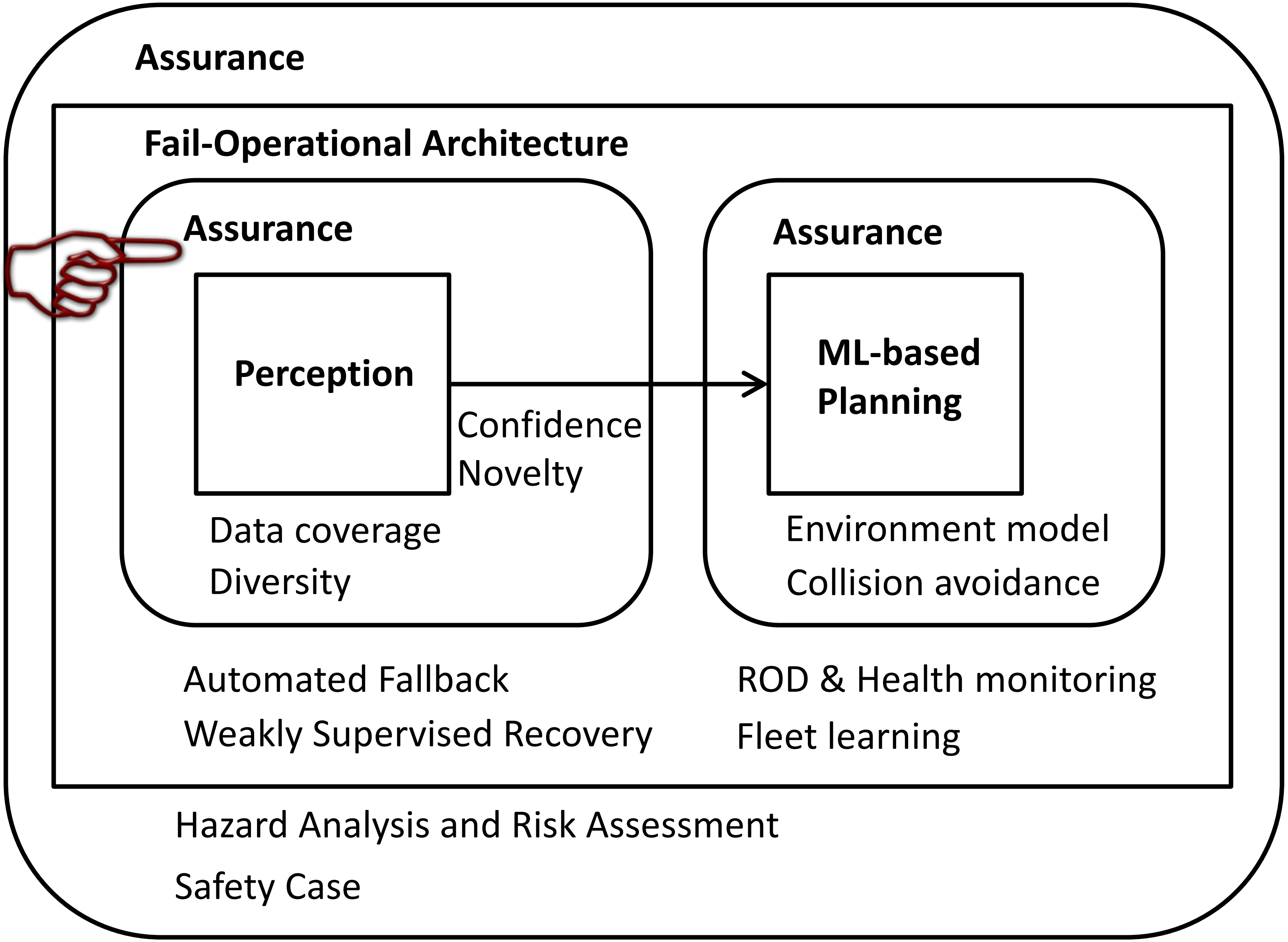
35 methods in Part 6 adapted
12 new methods specific to ML
Extensive literature review

https://uwaterloo.ca/wise-lab/projects/assuredai-safety-assurance-ai-based-automated-driving

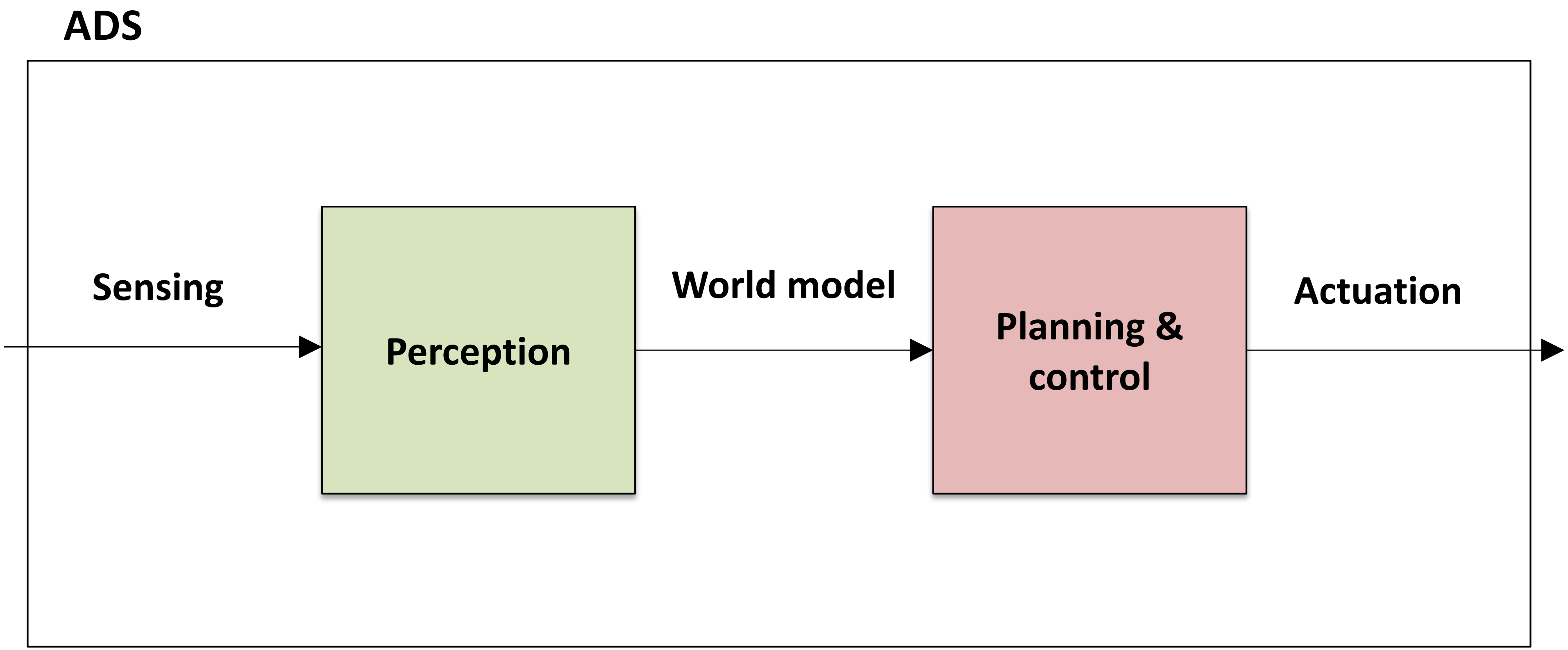
Process Extension Overview



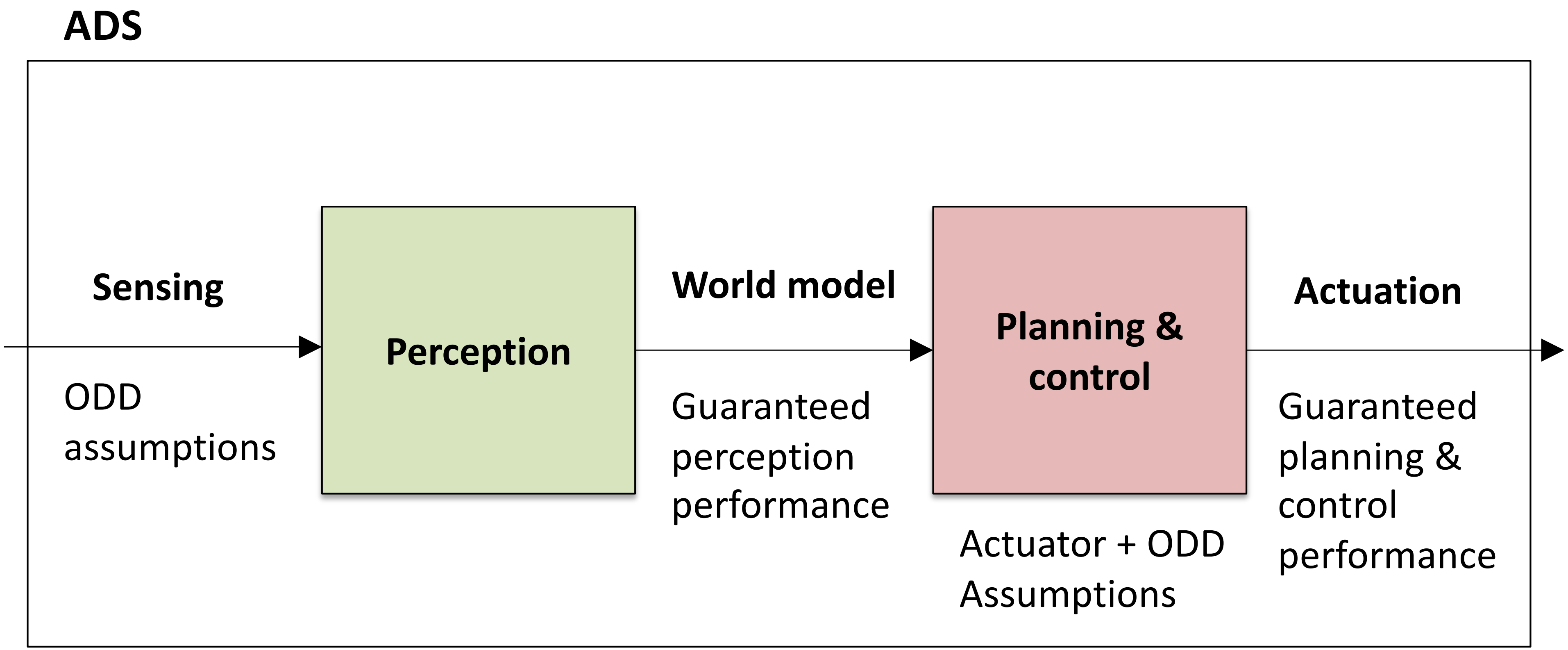
LAVA: Learned & Assured Vehicle Autonomy



Safety Argument Decomposition

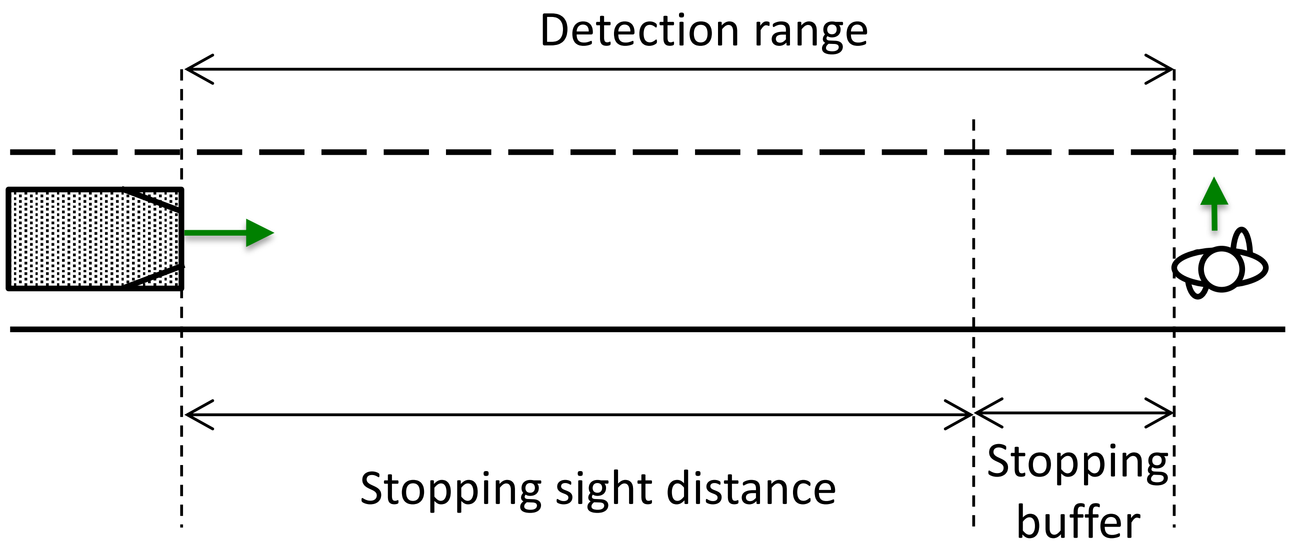


Safety Argument Decomposition



The following slides are based on Krzysztof Czarnecki and Rick Salay.
Towards a Framework to Manage Perceptual Uncertainty for Safe Automated Driving.
In WAISE, Västerås, Sweden, 2018
https://uwaterloo.ca/wise-lab/publications/towards-framework-manage-perceptual-uncertainty-safe

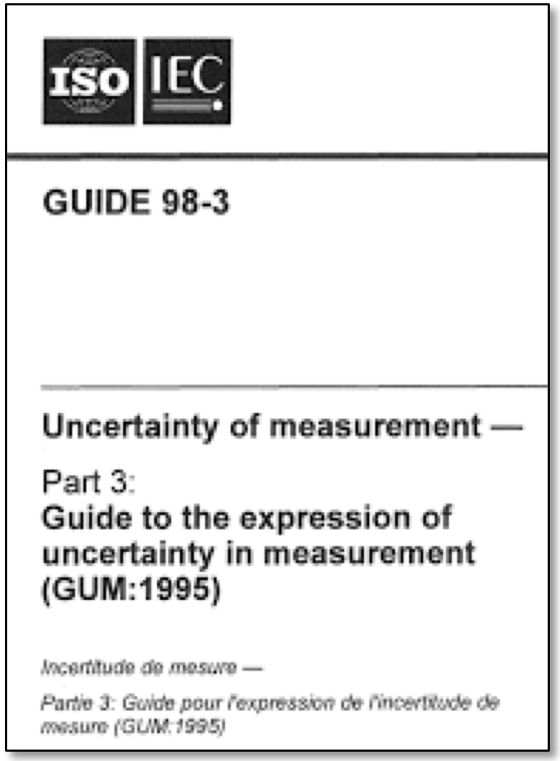
Sample Scenario-Dependent Perception-Performance Safety-Requirement Spec



Detect pedestrians on the roadway
within range of 10 m and with maximum perception-reaction delay of 0.5 s
with missed detection **probability** of 10^{-9} or less
with localization **uncertainty** of ± 0.5 m or better
within ODD conditions

Guide to the Expression of Uncertainty in Measurement (GUM)

- True accuracy unknowable
  - Accuracy in ML wrt. test set only
- Must estimate uncertainty


ISO IEC

GUIDE 98-3

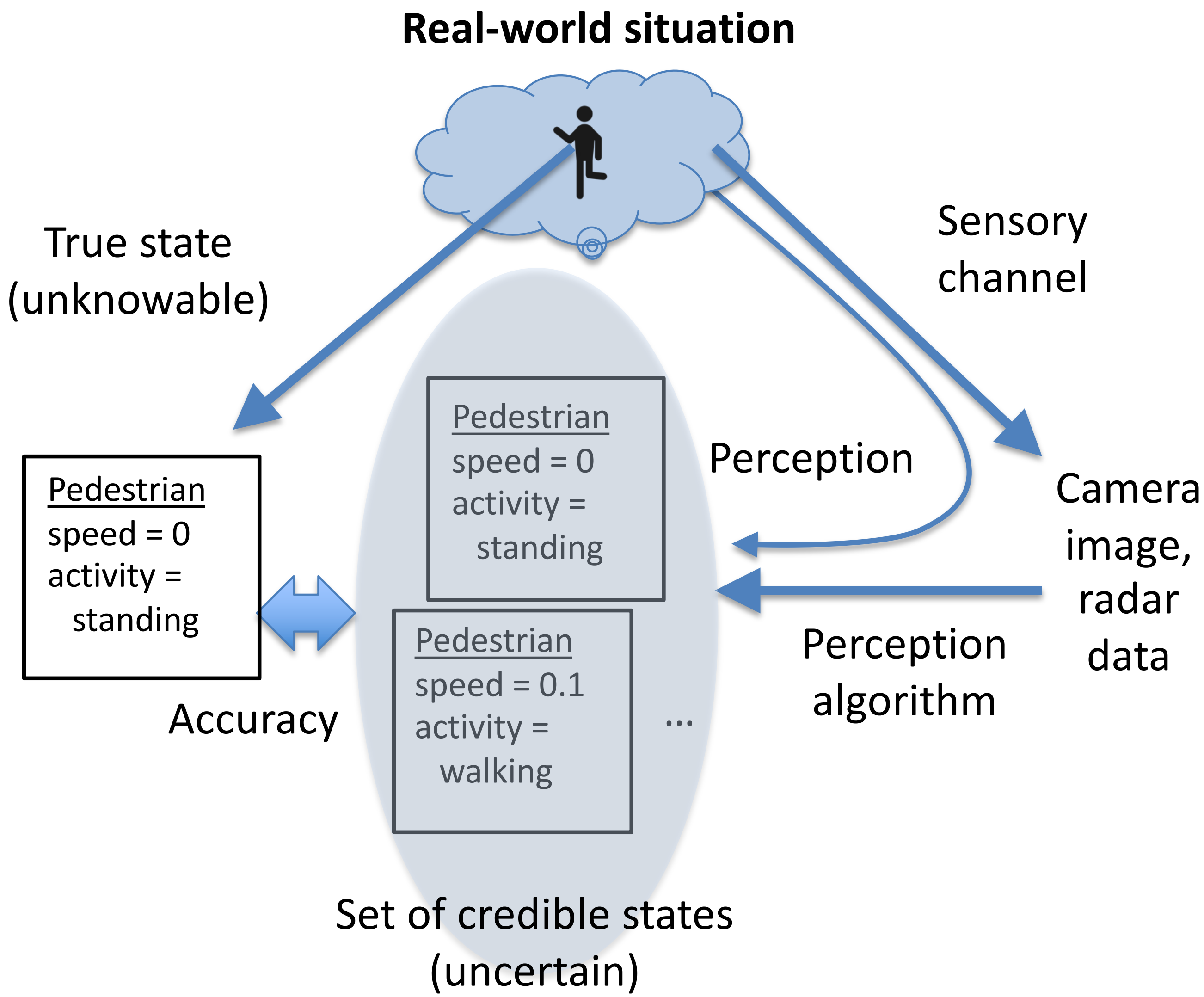
Uncertainty of measurement —

Part 3:
Guide to the expression of uncertainty in measurement (GUM:1995)

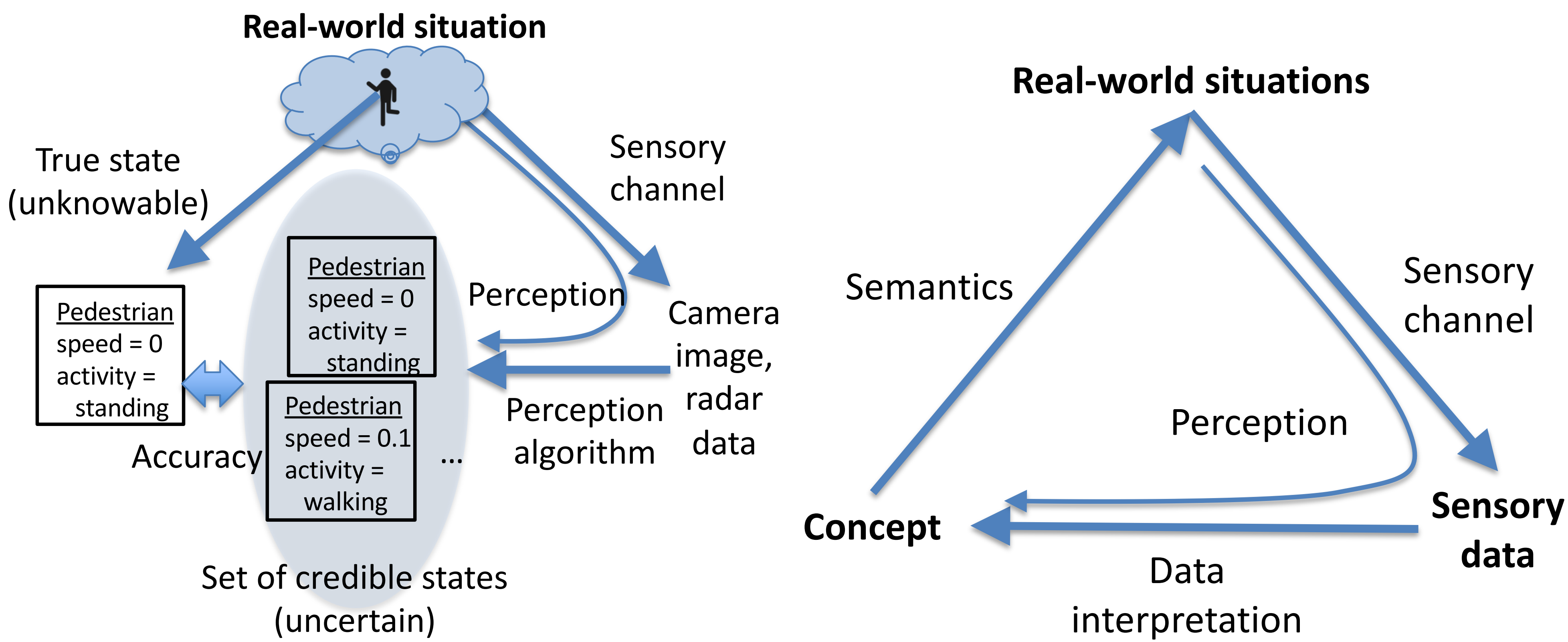
Incertitude de mesure —

Partie 3: Guide pour l'expression de l'incertitude de mesure (GUM:1995)

Perception Triangle (Instance-Level)

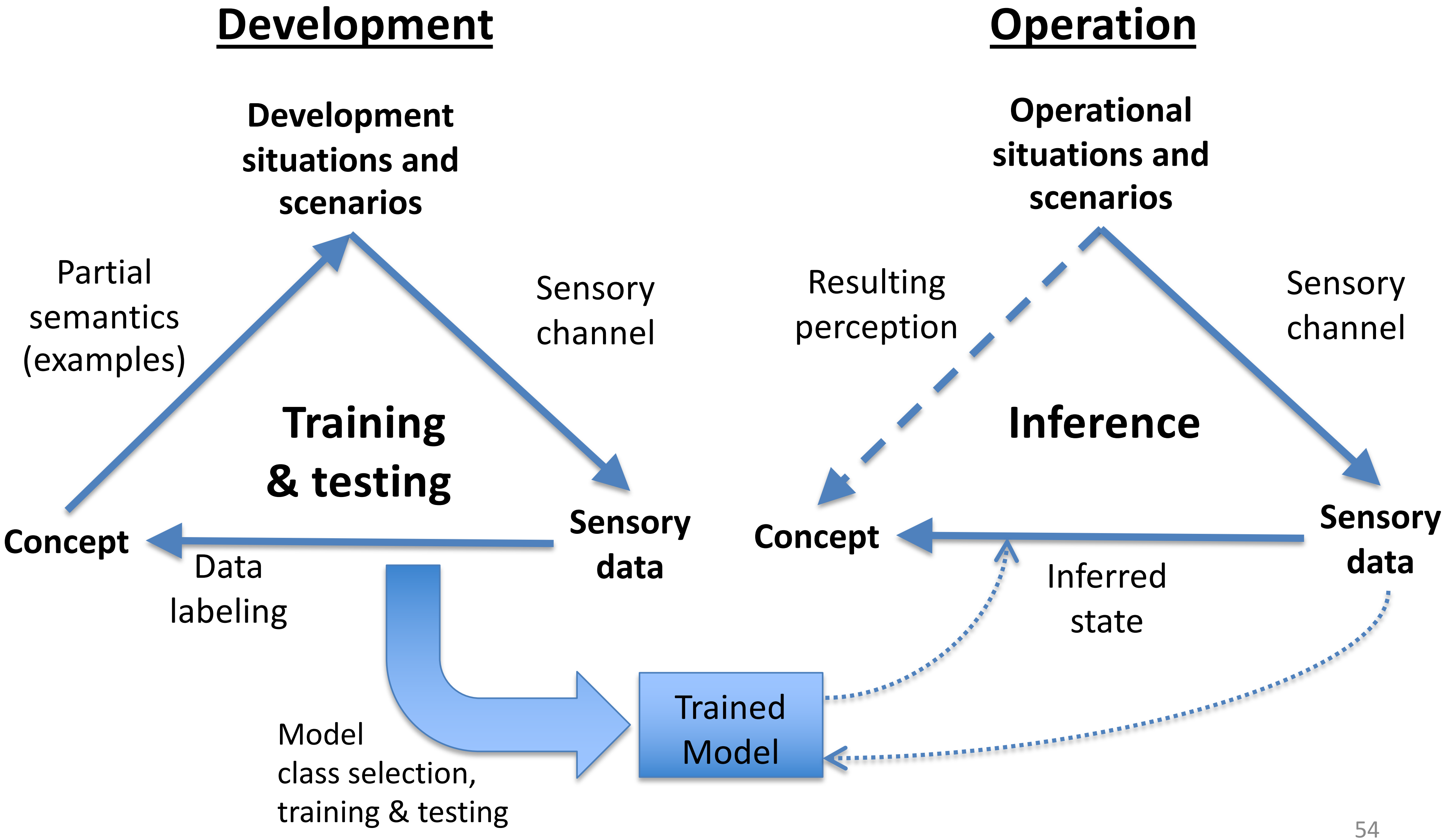


Perceptual Triangle

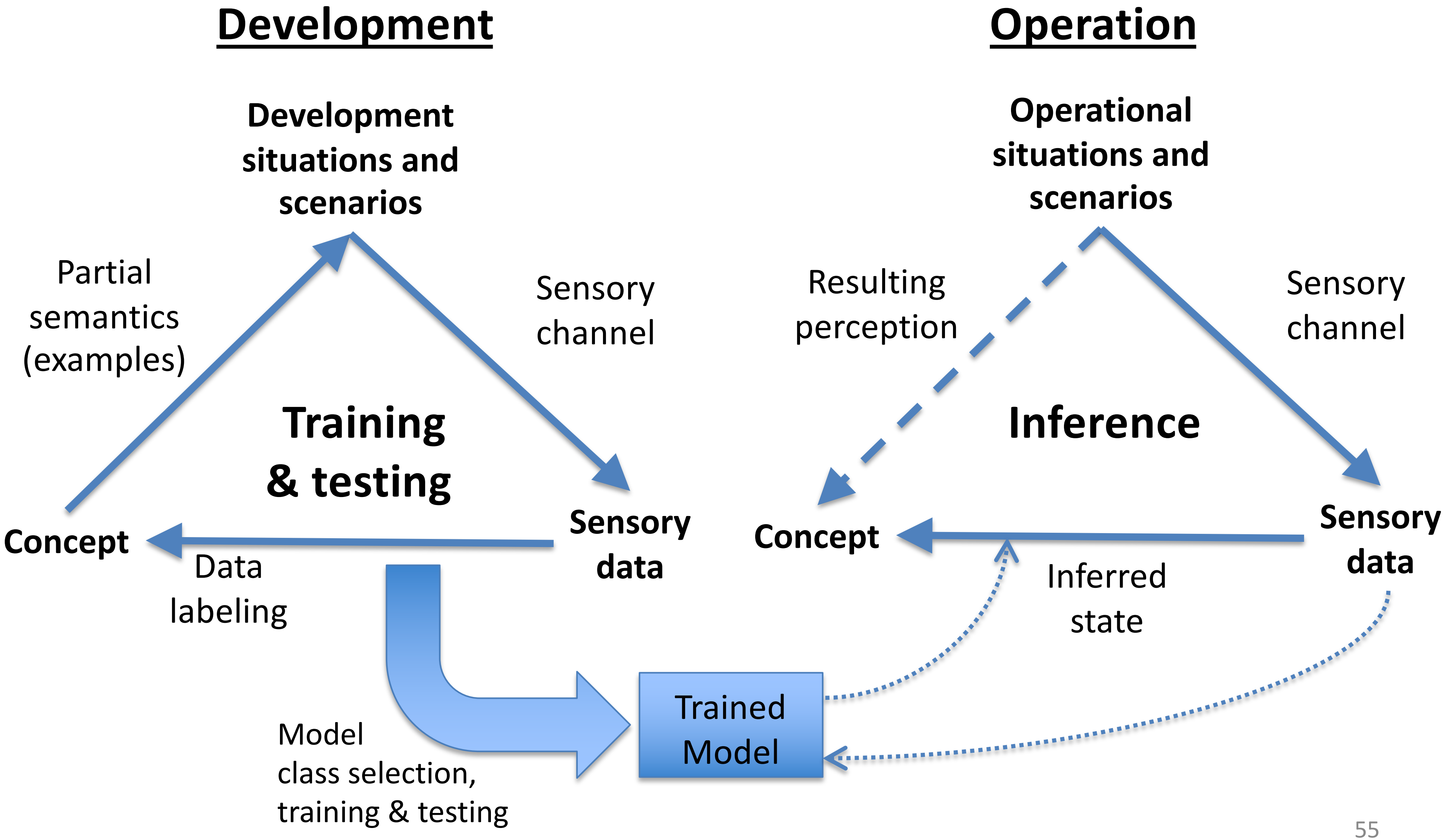


Instance-level **Domain-level (generic)**

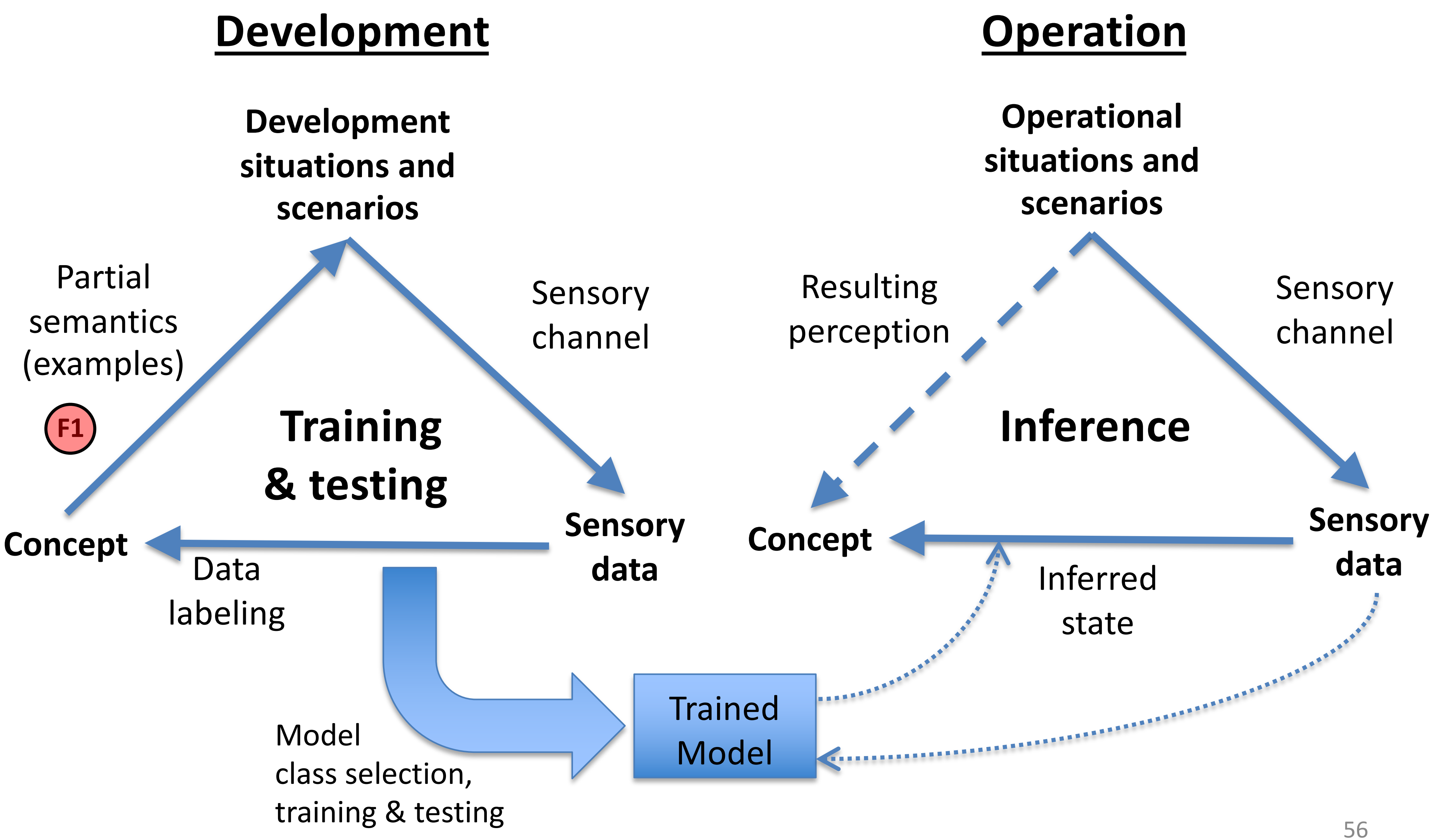
Perceptual Triangle When Using Supervised ML



Factors Influencing Uncertainty



F1: Conceptual Uncertainty



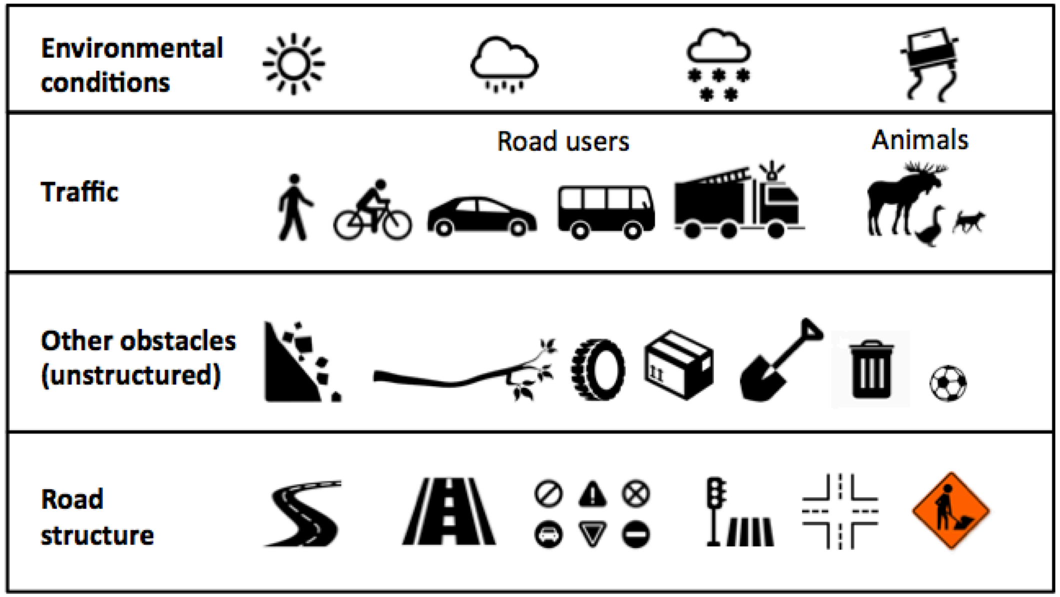
F1: Conceptual Uncertainty
Pedestrian or Cyclist?





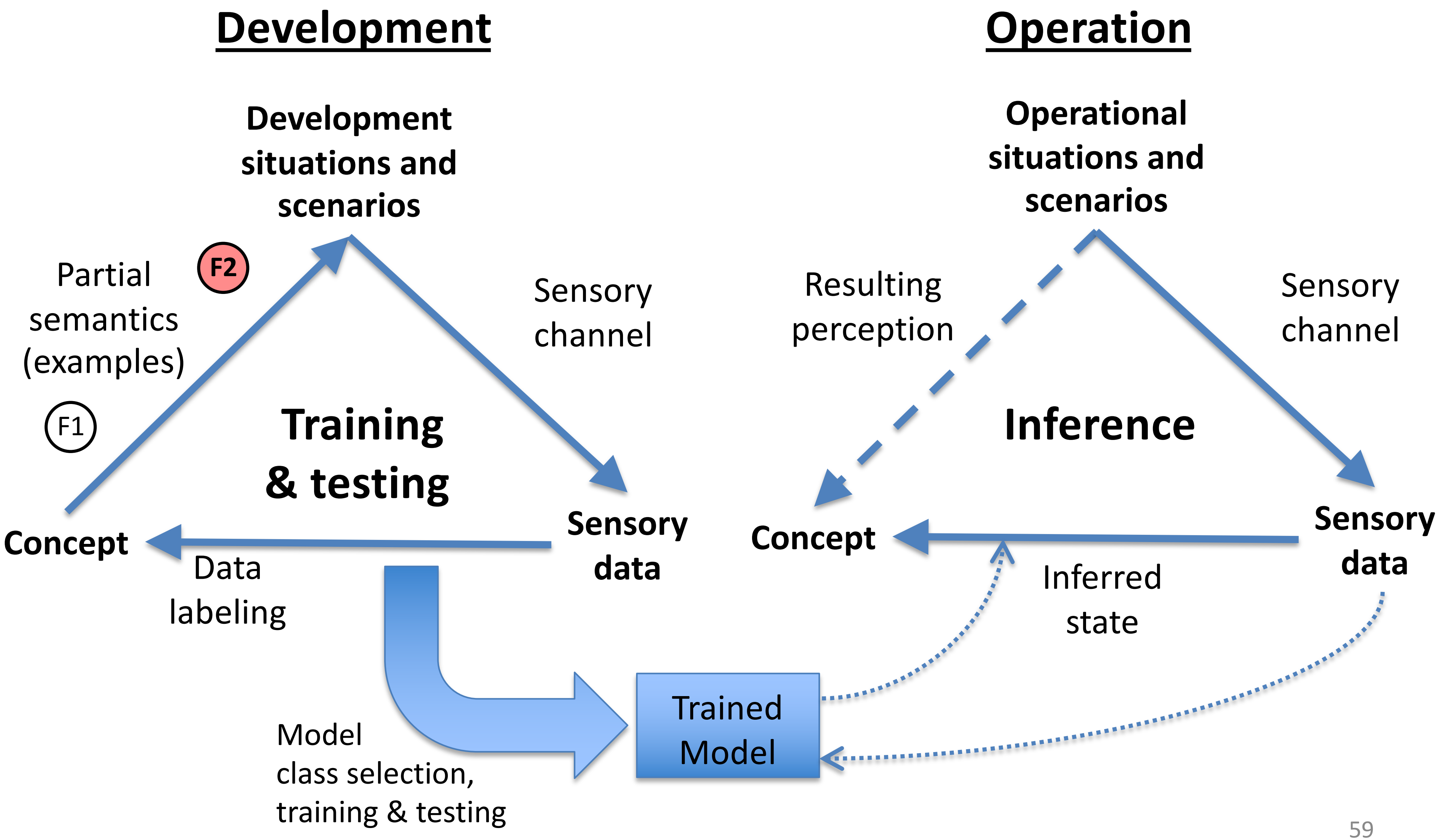
F1: Conceptual Uncertainty

- Assessed by expert review or labeling disagreement
- Reduced by developing standard ontologies
  - E.g., WISE Drive Ontology

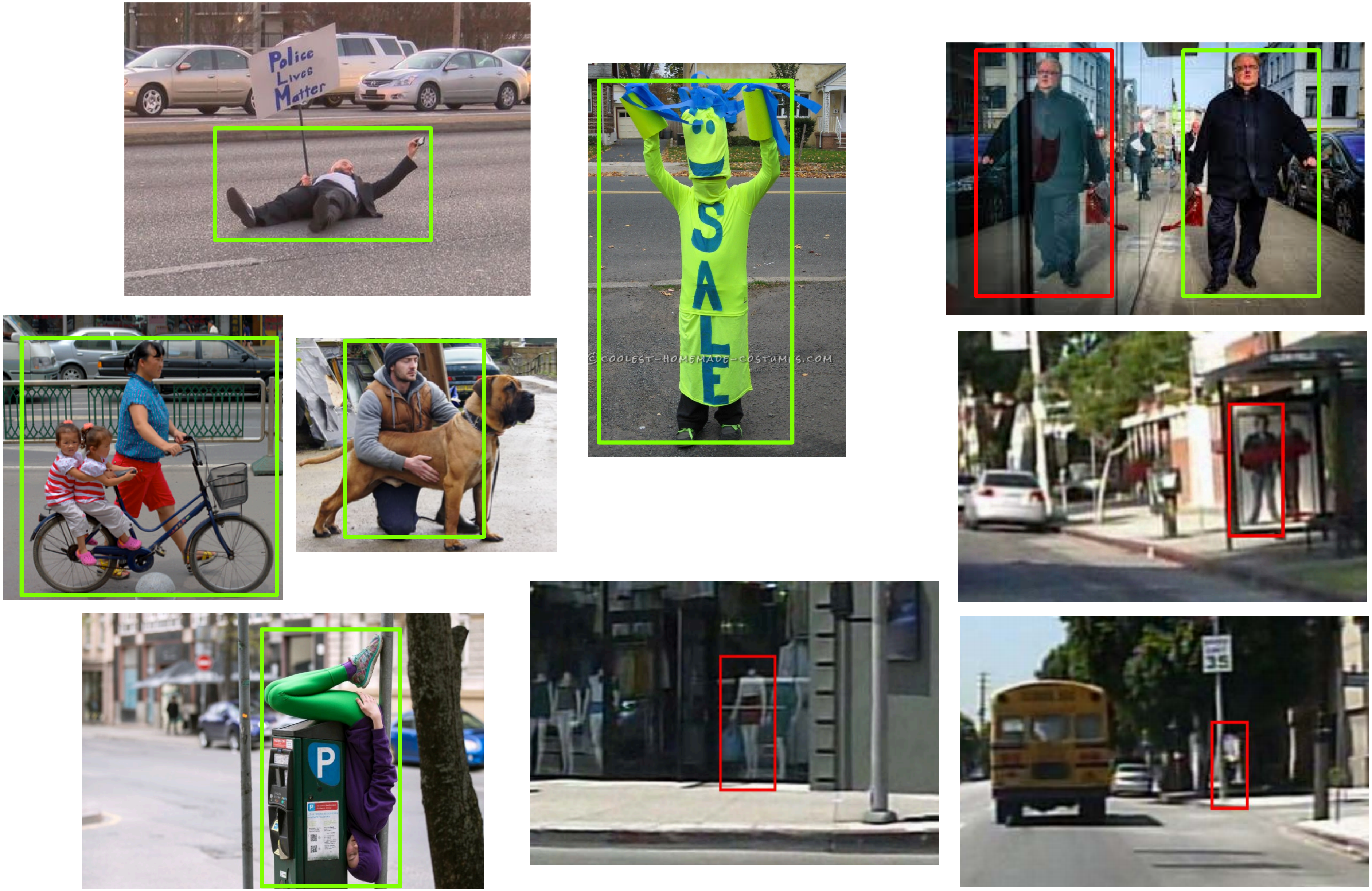


https://uwaterloo.ca/wise-lab/projects/wise-drive-requirements-analysis-framework-automated-driving

F2: Development Scenario Coverage

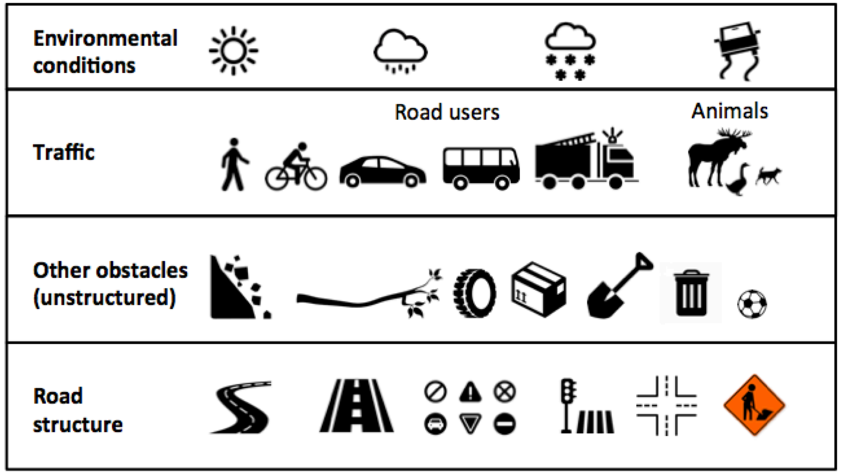


F2: Development Scenario Coverage



F2: Development Scenario Coverage

- Assessed with respect to ontologies and field validation targets
  - Must include positive/negative and near-hit/near-miss examples



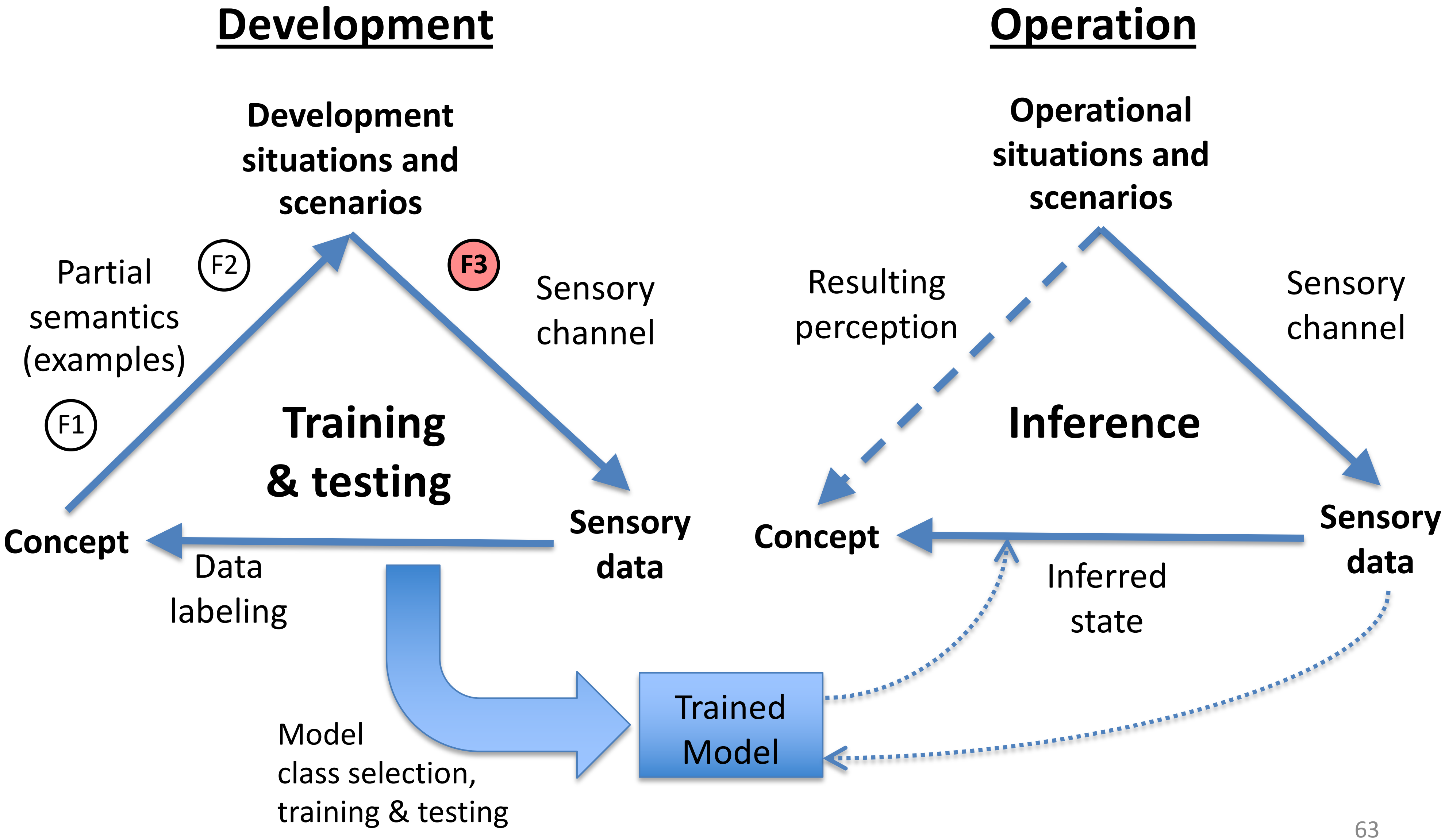
- Challenge: how much data is enough?

Active Learning

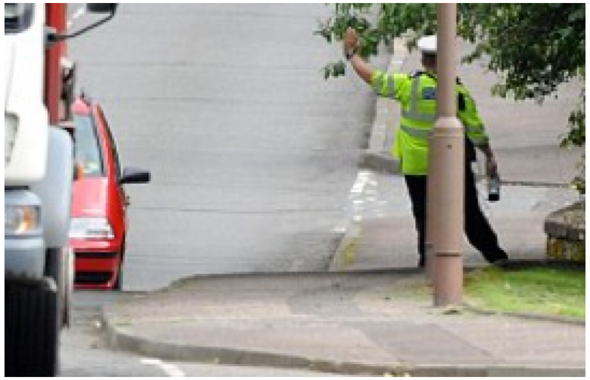
Data selection criteria

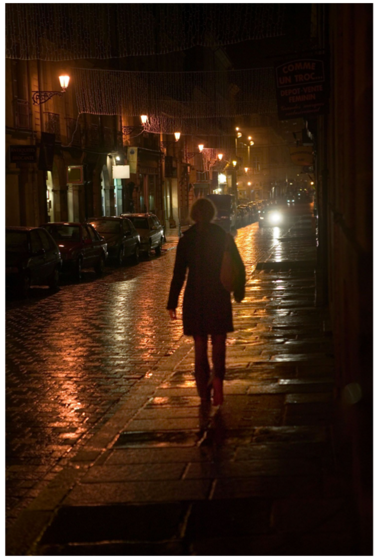
1. Uncertainty
2. Coverage & diversity
3. Collection & labeling cost
4. Risk profile

F3: Scene Uncertainty



F3: Scene Uncertainty

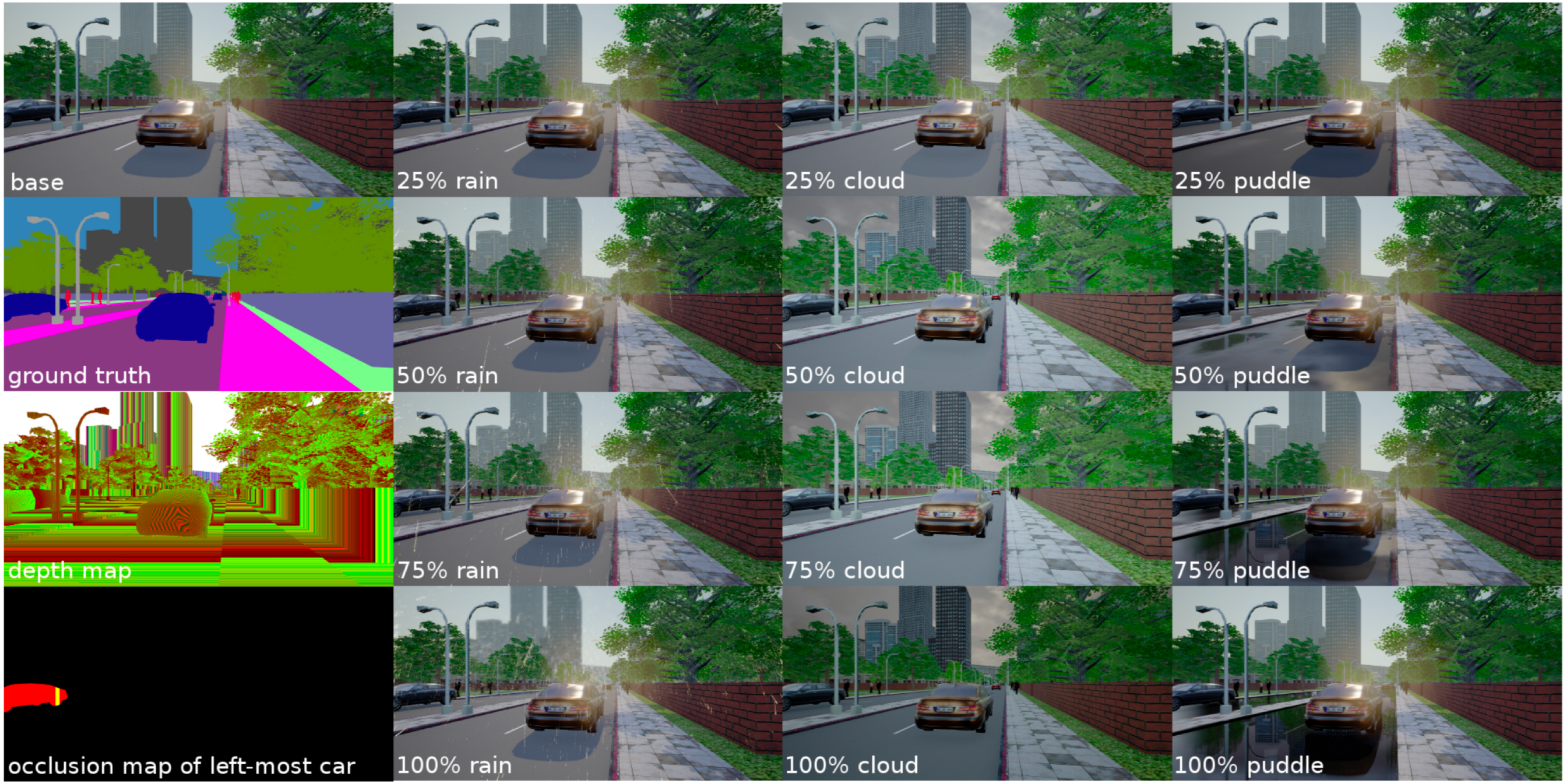




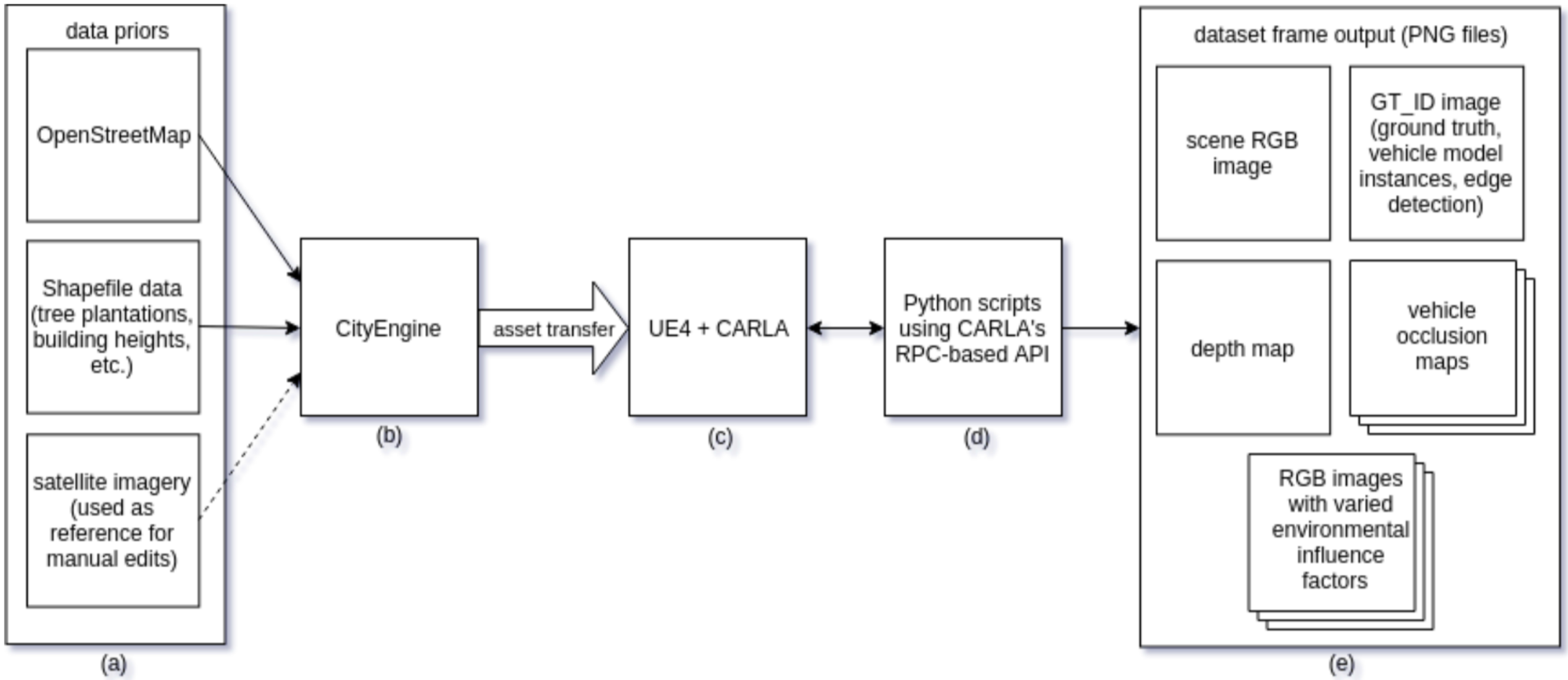
F3: Scene Uncertainty

- Surrogate measures
  - range, scale, occlusion level, atmospheric visibility, illumination, clutter and crowding level
- May compare test set accuracy and output confidence with these measures
- Also part of development data set coverage

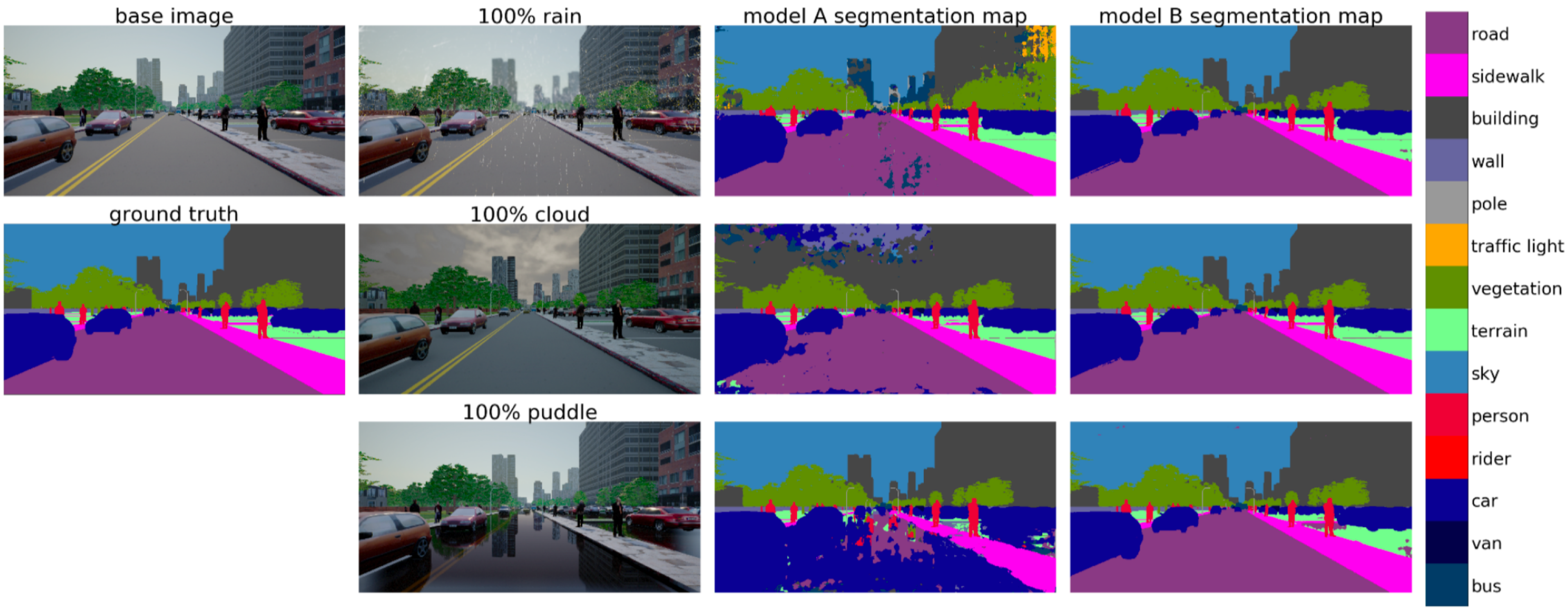
Synthetic Dataset to Study Scene Influence Factors

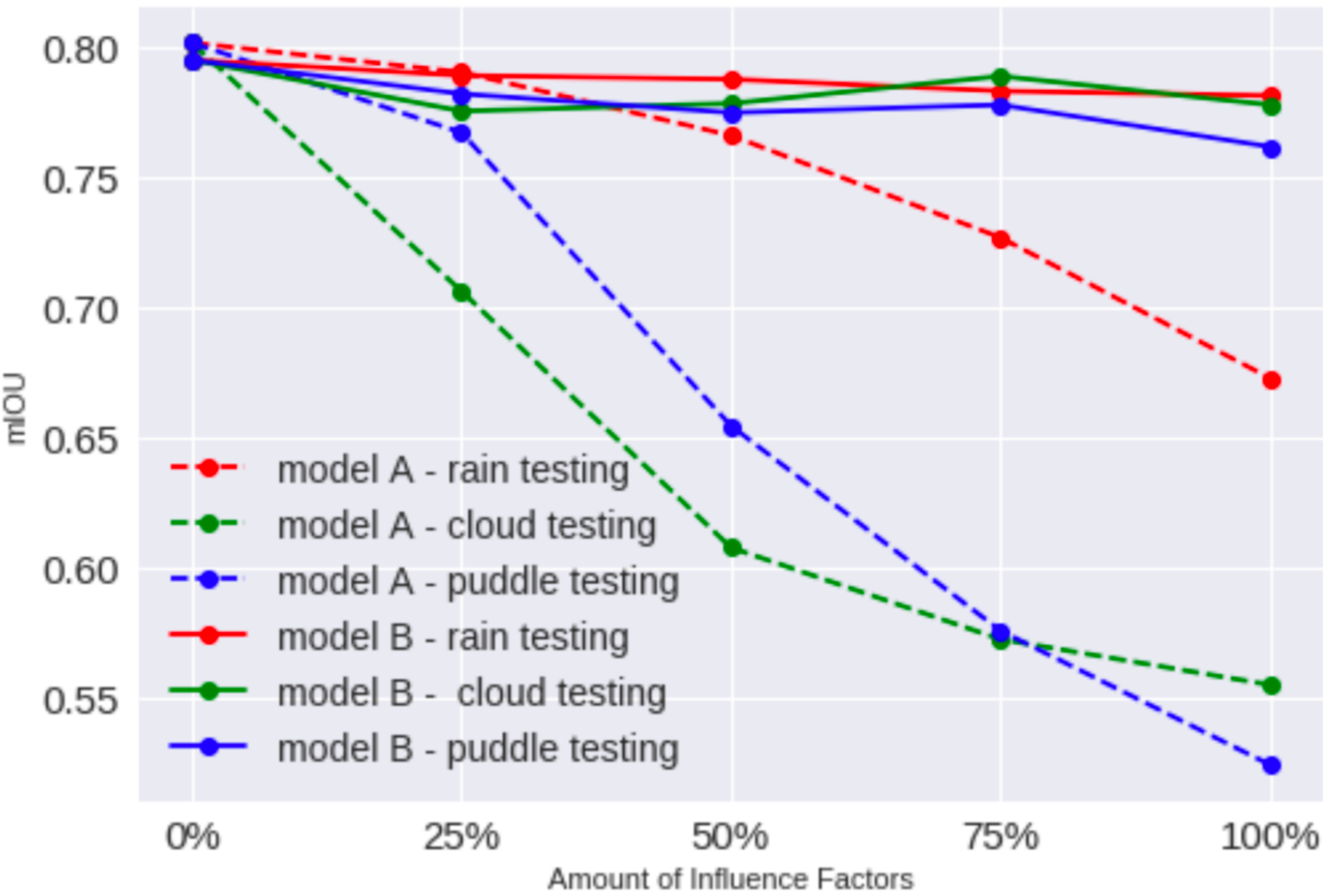


Data Generation Pipeline



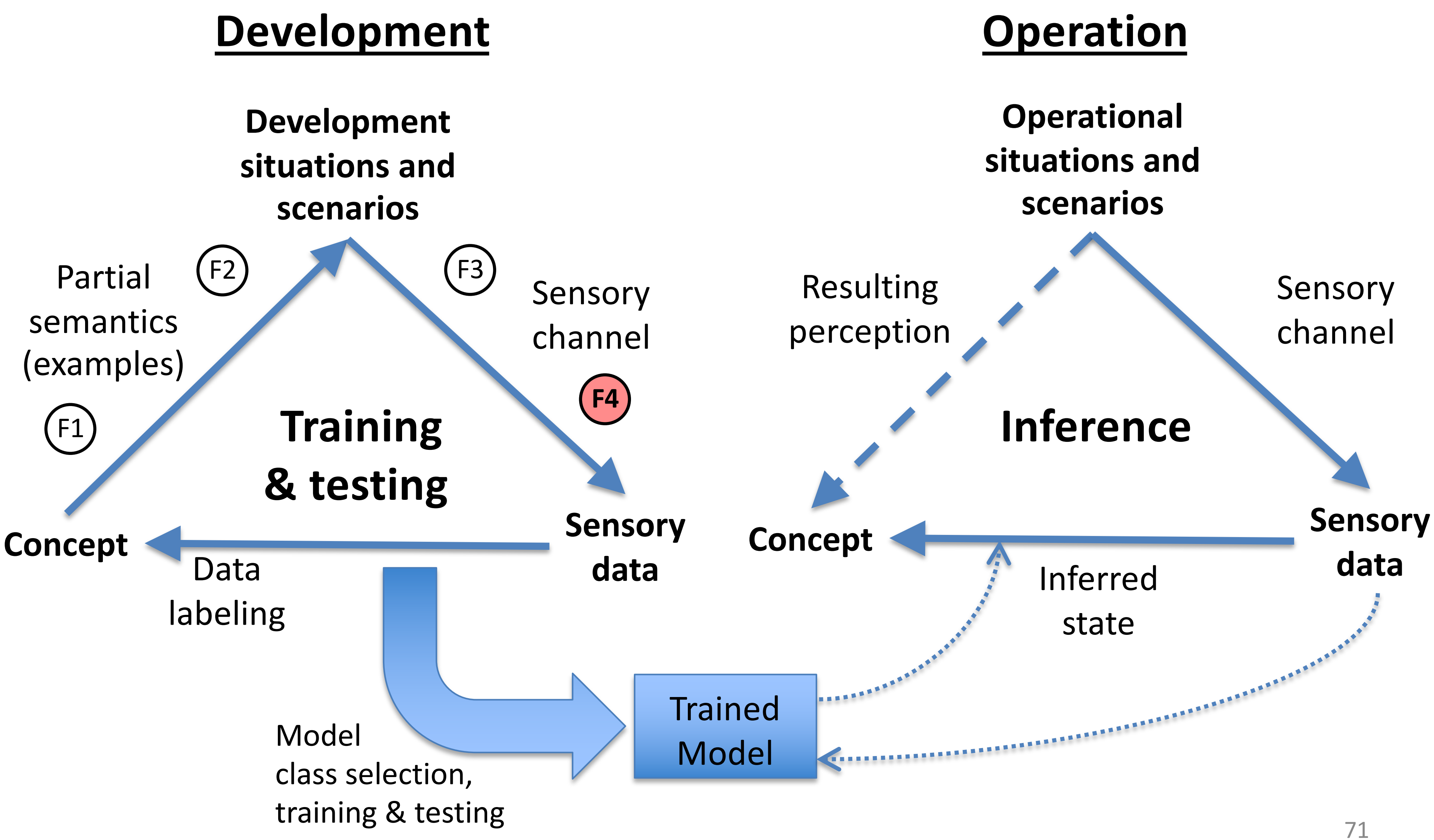
Scene Influence Factors -> Accuracy



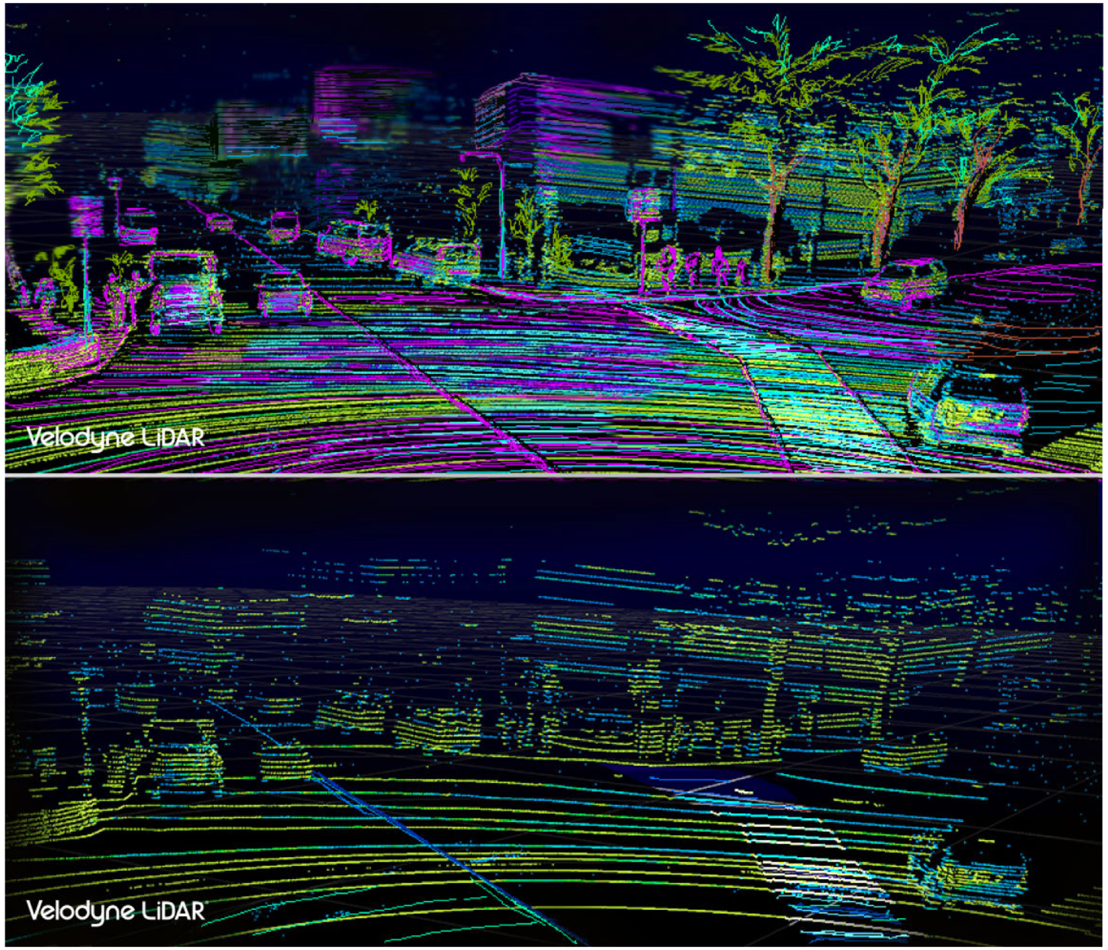

0.80
0.75
0.70
0.65
0.60
0.55
mIOU
model A - rain testing
model A - cloud testing
model A - puddle testing
model B - rain testing
model B - cloud testing
model B - puddle testing
0%
25%
50%
75%
100%
Amount of Influence Factors


0.74
0.72
0.70
0.68
0.66
0.64
0.62
mIOU
0%
25%
50%
75%
100%
Rain Variations
3000 clean images
3000 clean and 100 rainy images
3000 clean and 500 rainy images
3000 clean and 1000 rainy images

F4: Sensor Properties



F4: Sensor Properties



Daylight White Balance

Cloudy White Balance

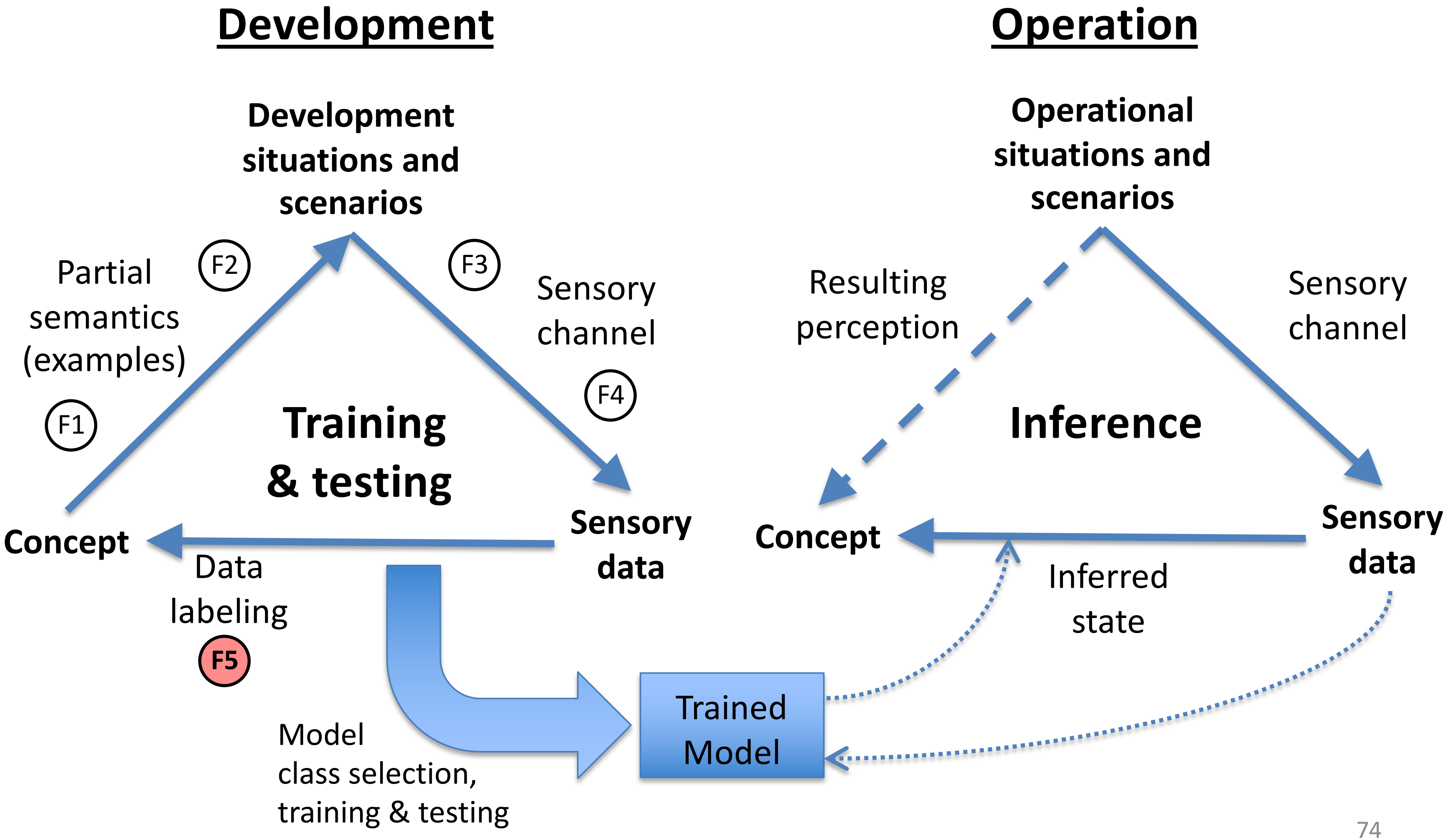
Shade White Balance

Tungsten White Balance

F4: Sensor Properties

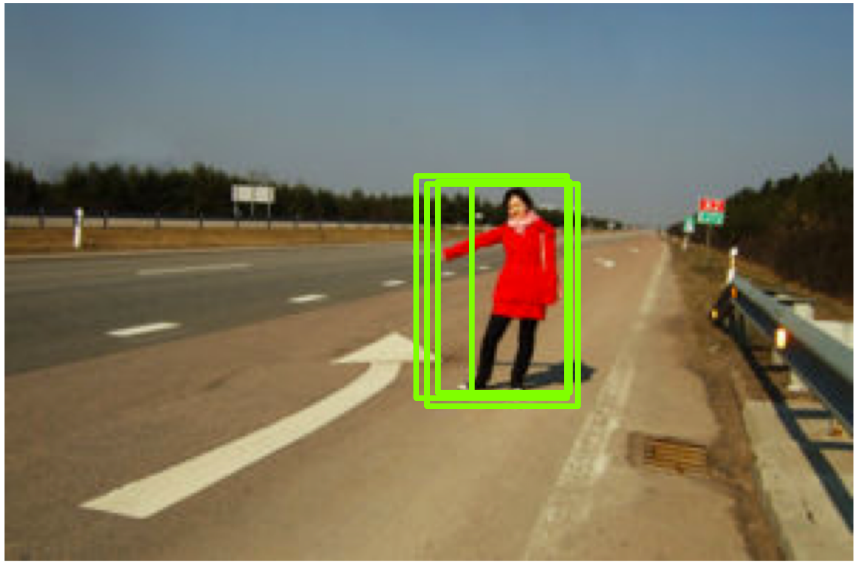
- Mature engineering discipline
  - Determining sensor properties to capture sufficient information
  - Mode, range, resolution, sensitivity, placement, etc.
- However, interaction between ML algorithms and sensor properties must be assessed
  - E.g., how effective is ML is ignoring sensor noise or artifacts?

F5: Label Uncertainty

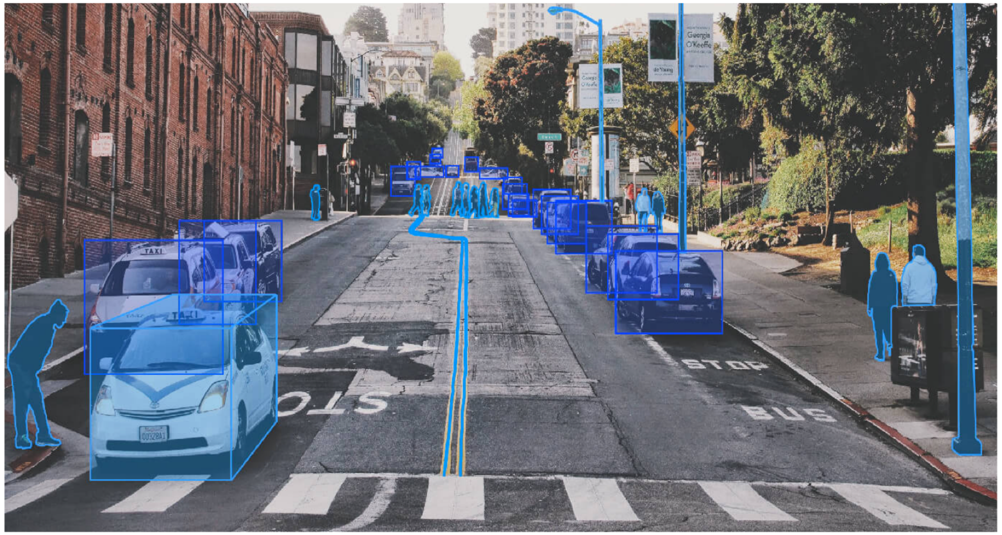


F5: Label Uncertainty

Class: cyclist vs. pedestrian



Bounding box placement uncertainty

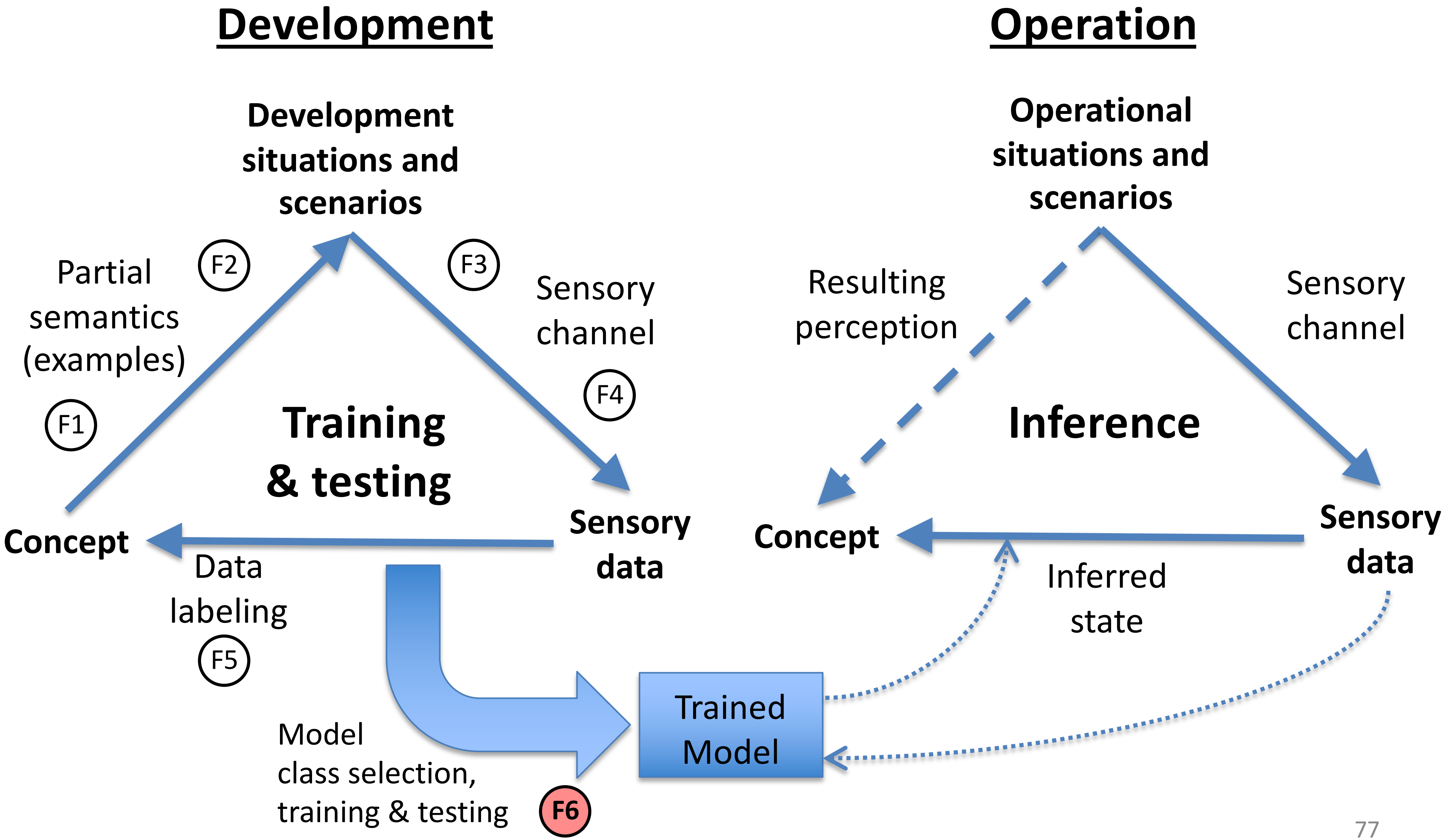


3D bounding box placement is challenging

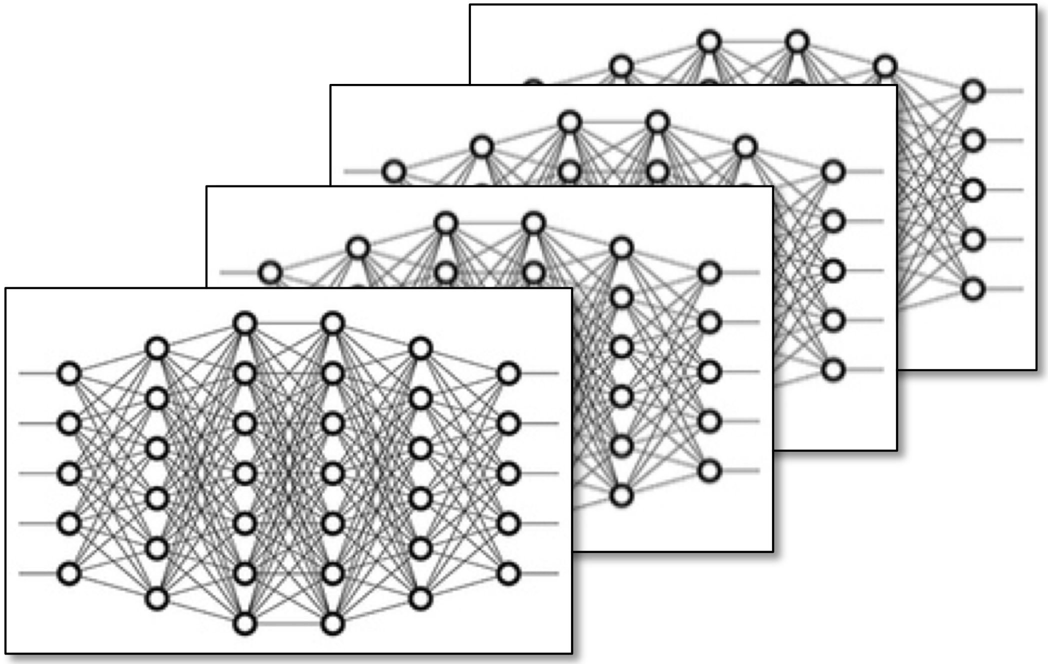
F5: Label Uncertainty

- Assessed by expert review and labeler disagreement
  - Existing research on determining number of labelers in crowd sourcing
  - E.g., may need as many as 6 independent votes
- Reduction measures
  - Conceptual clarity (F1)
  - Quality control
    - Clear instructions, training, verification, etc.
    - Bread and butter of labeling companies

F6: Model Uncertainty



F6: Model Uncertainty



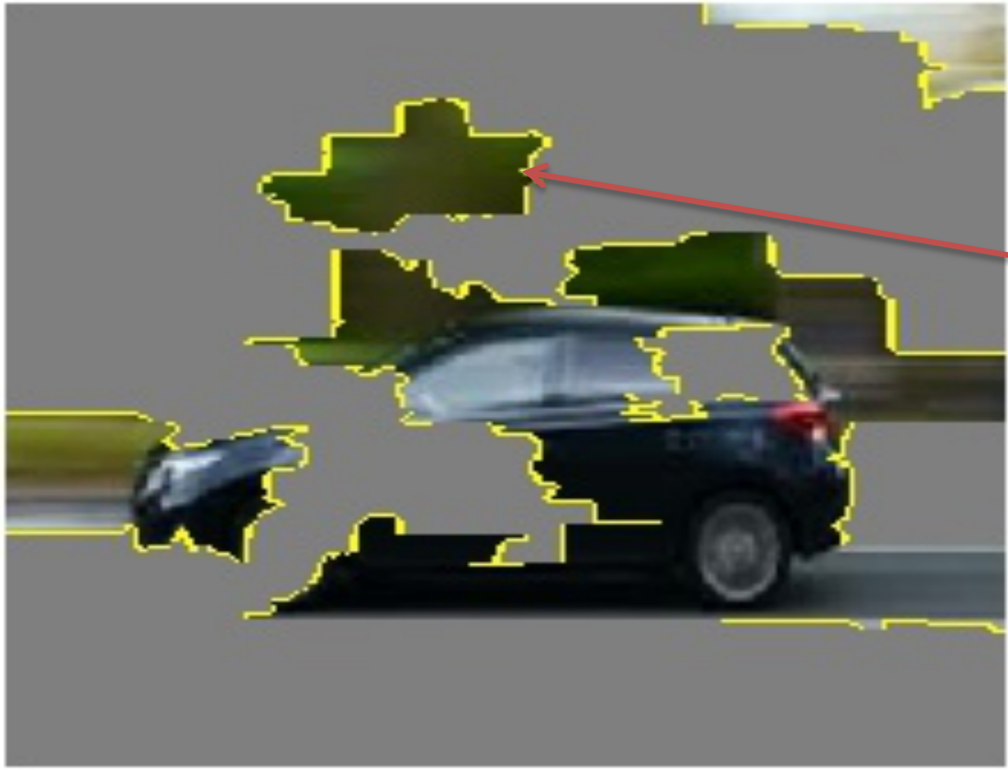
What model was learned in training?
What decisions will it make in operation?

F6: Model Uncertainty

1. Explanation methods help validate features
2. Robustness measures help assess risk of misclassification
3. Bayesian deep learning can help assess model uncertainty

Deep Learning and Explanations

Passenger car



The explanation shows
that a tree
contributed
to the classification
decision
(method: LIME)

The top 15 features (superpixels) used to classify corresponding input image as a car by an Inception network trained on ImageNet

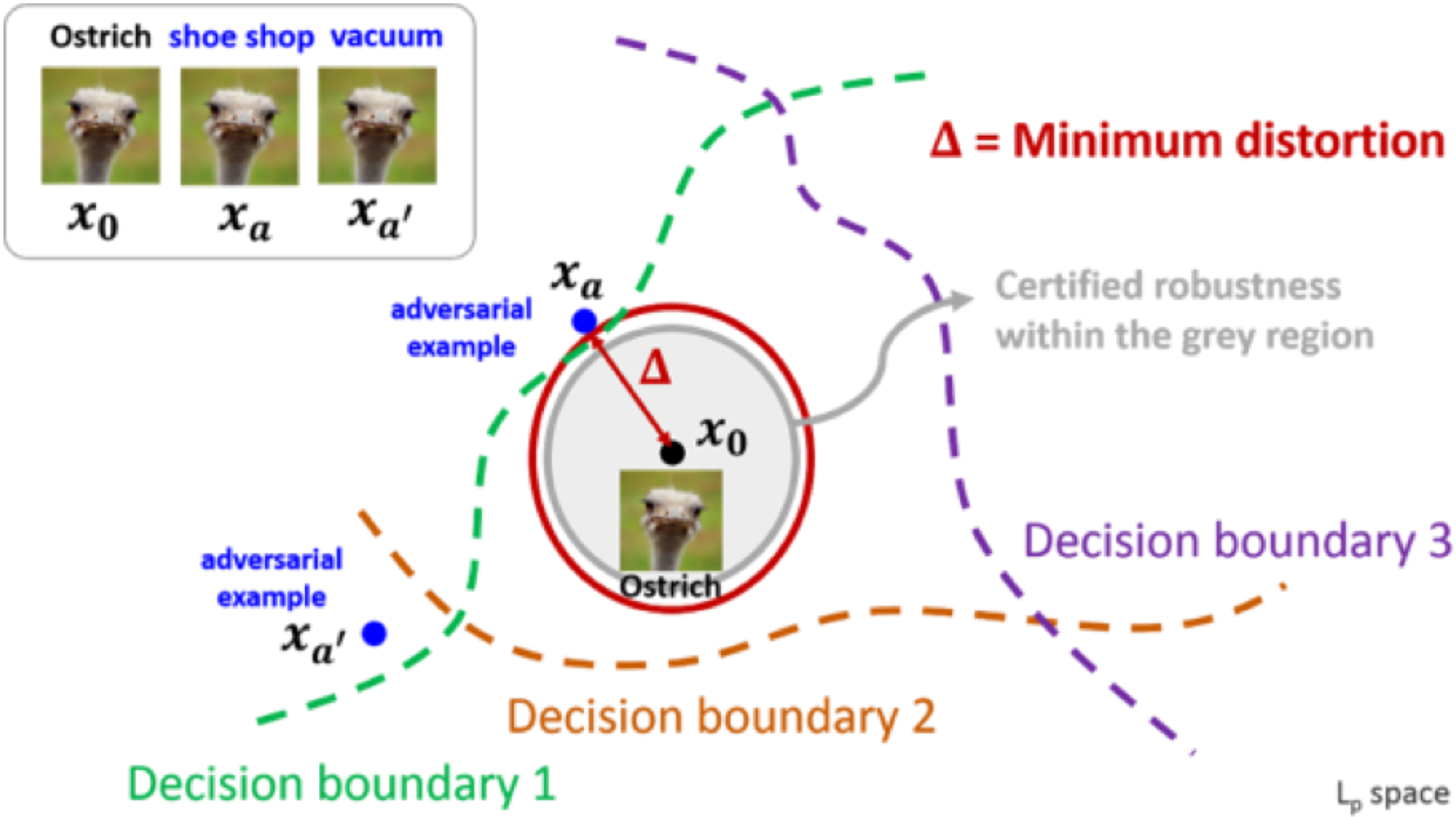
(see LIME at https://github.com/marcotcr/lime)

Adversarial Stickers

Misclassified as speed signs

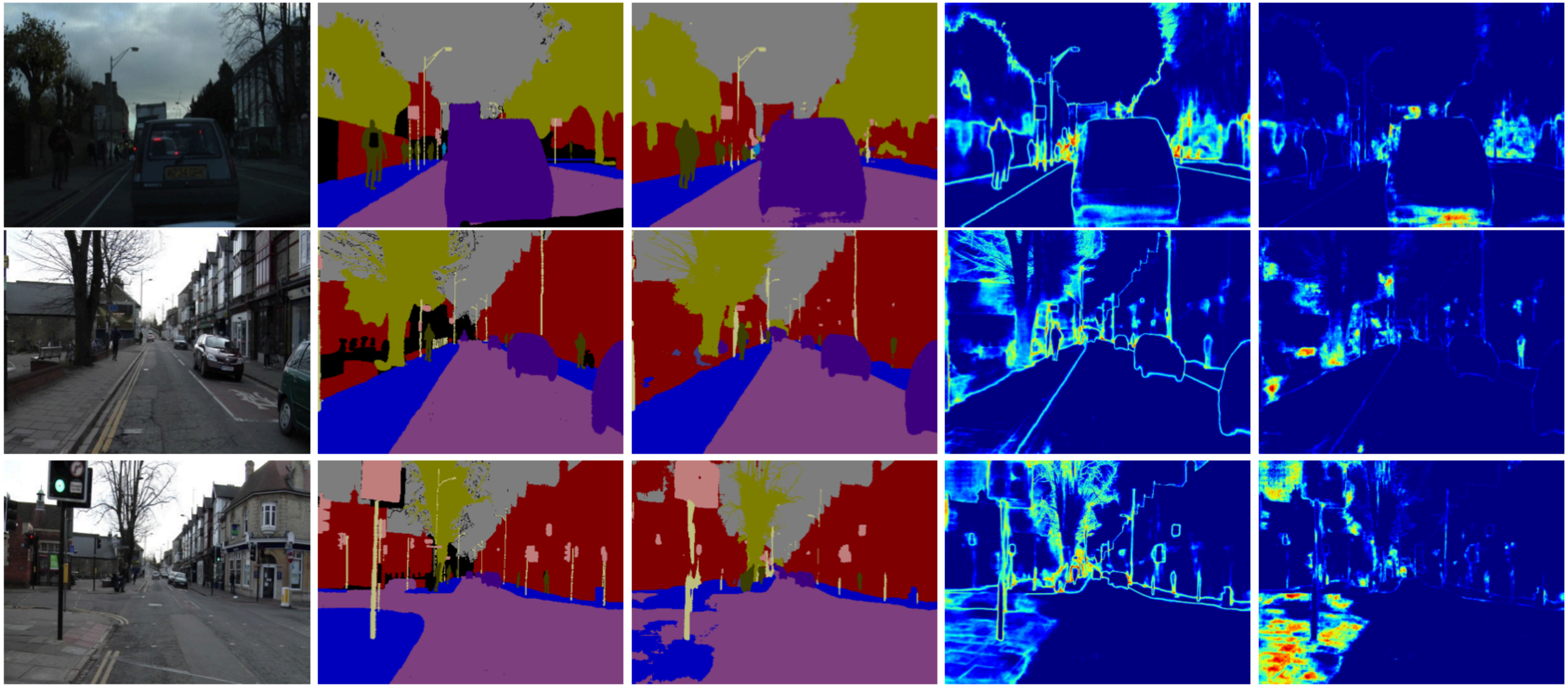
Evtimov et al.

Robustness Measures



CLEVER approach by IBM

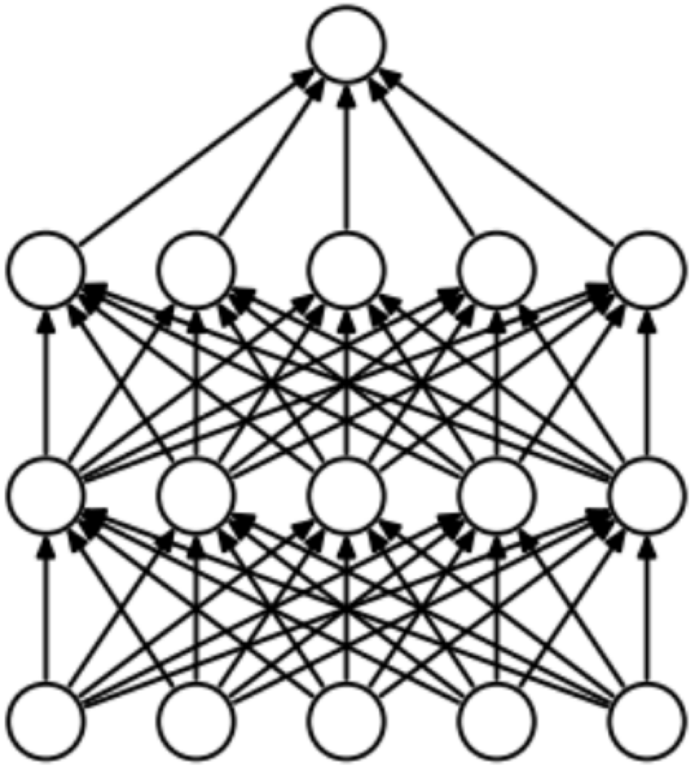
Aleatoric and Epistemic Uncertainty



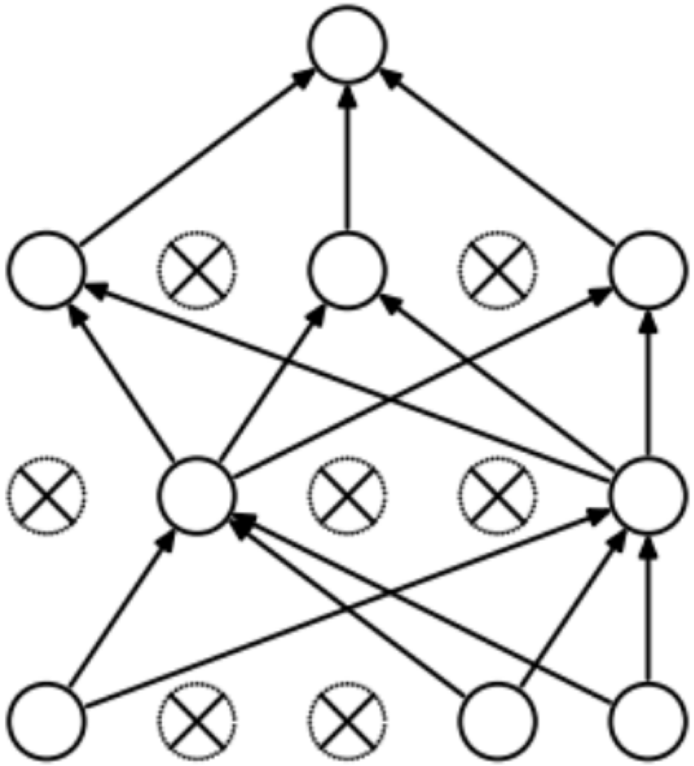
(a) Input Image (b) Ground Truth (c) Semantic Segmentation (d) Aleatoric Uncertainty (e) Epistemic Uncertainty

Yarin Gal, et al., https://arxiv.org/abs/1703.04977

Dropout

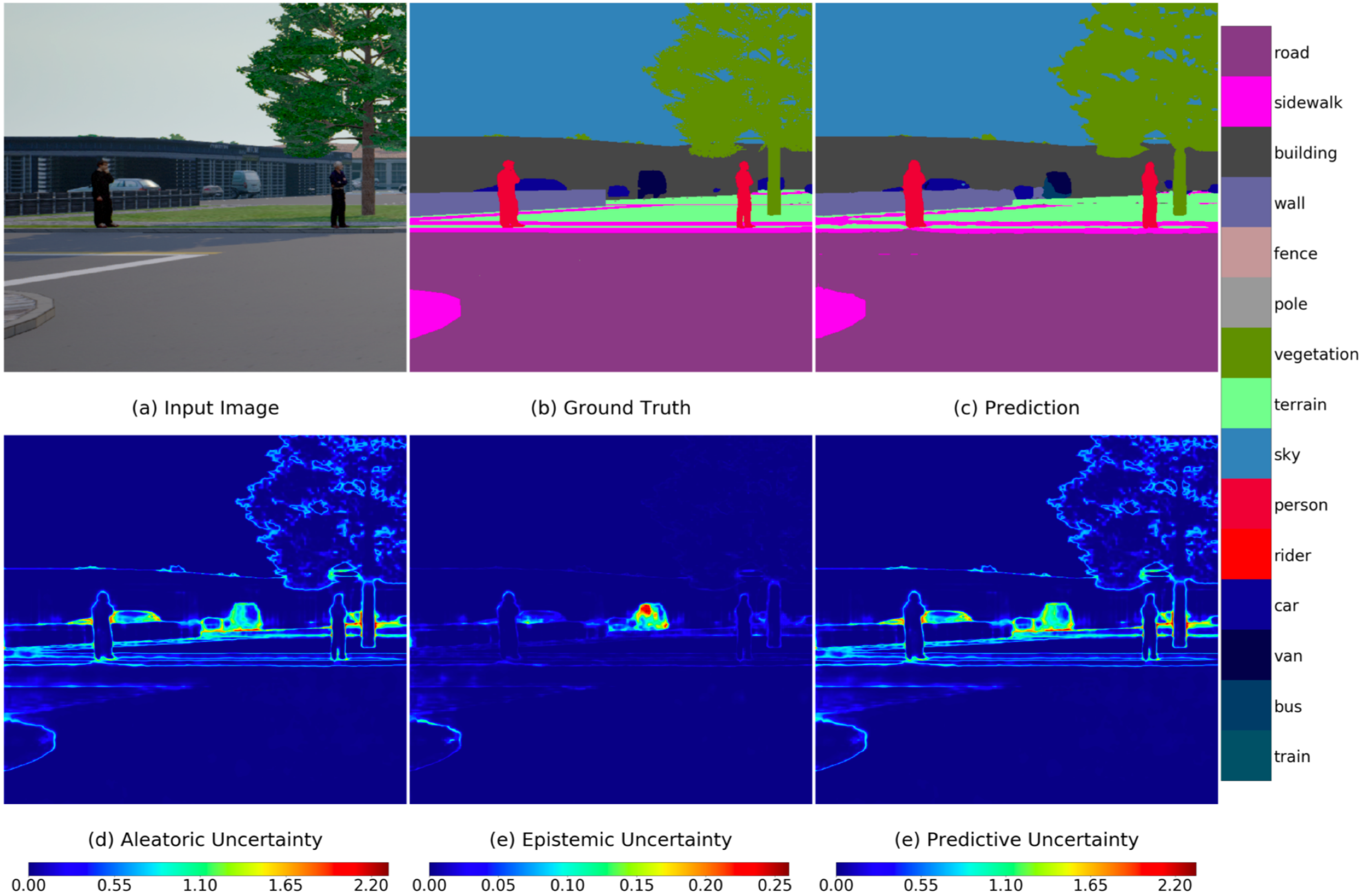


(a) Standard Neural Net

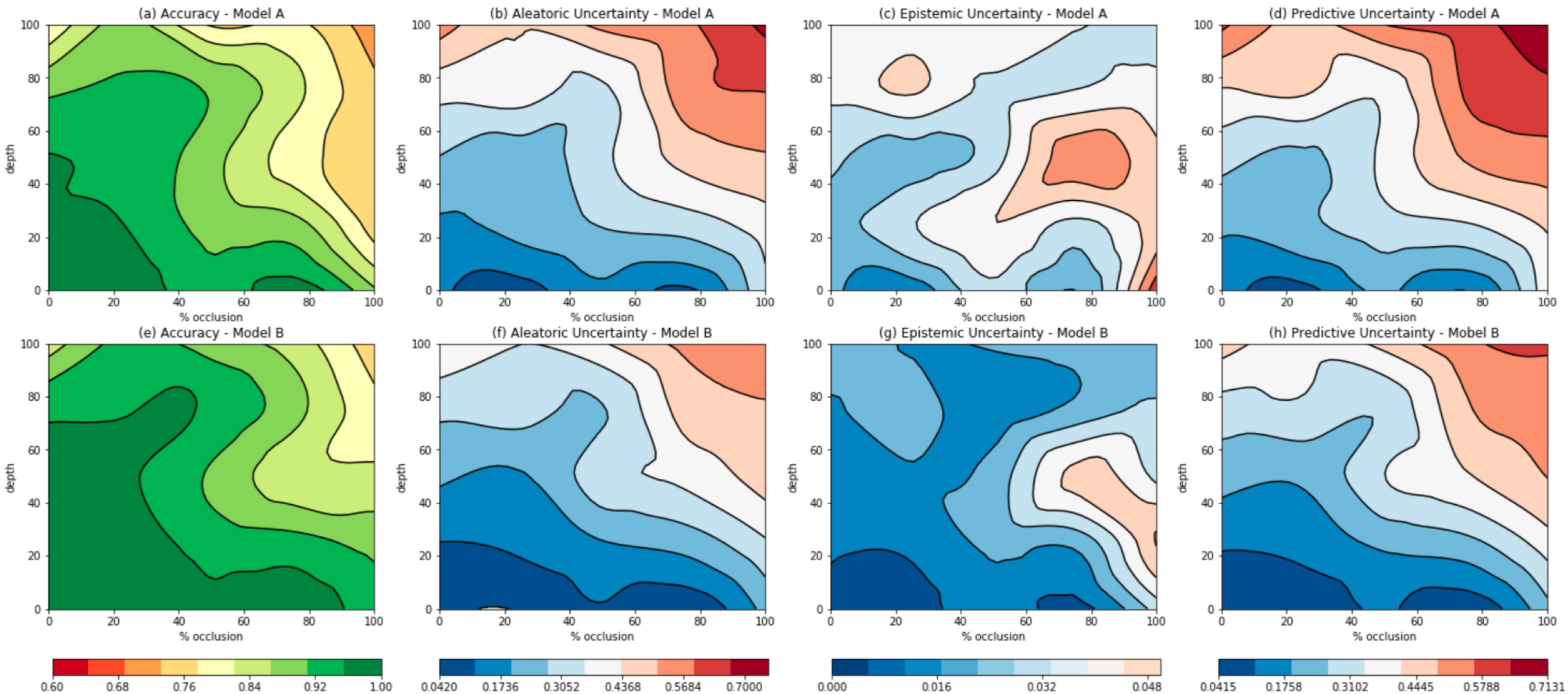


(b) After applying dropout.

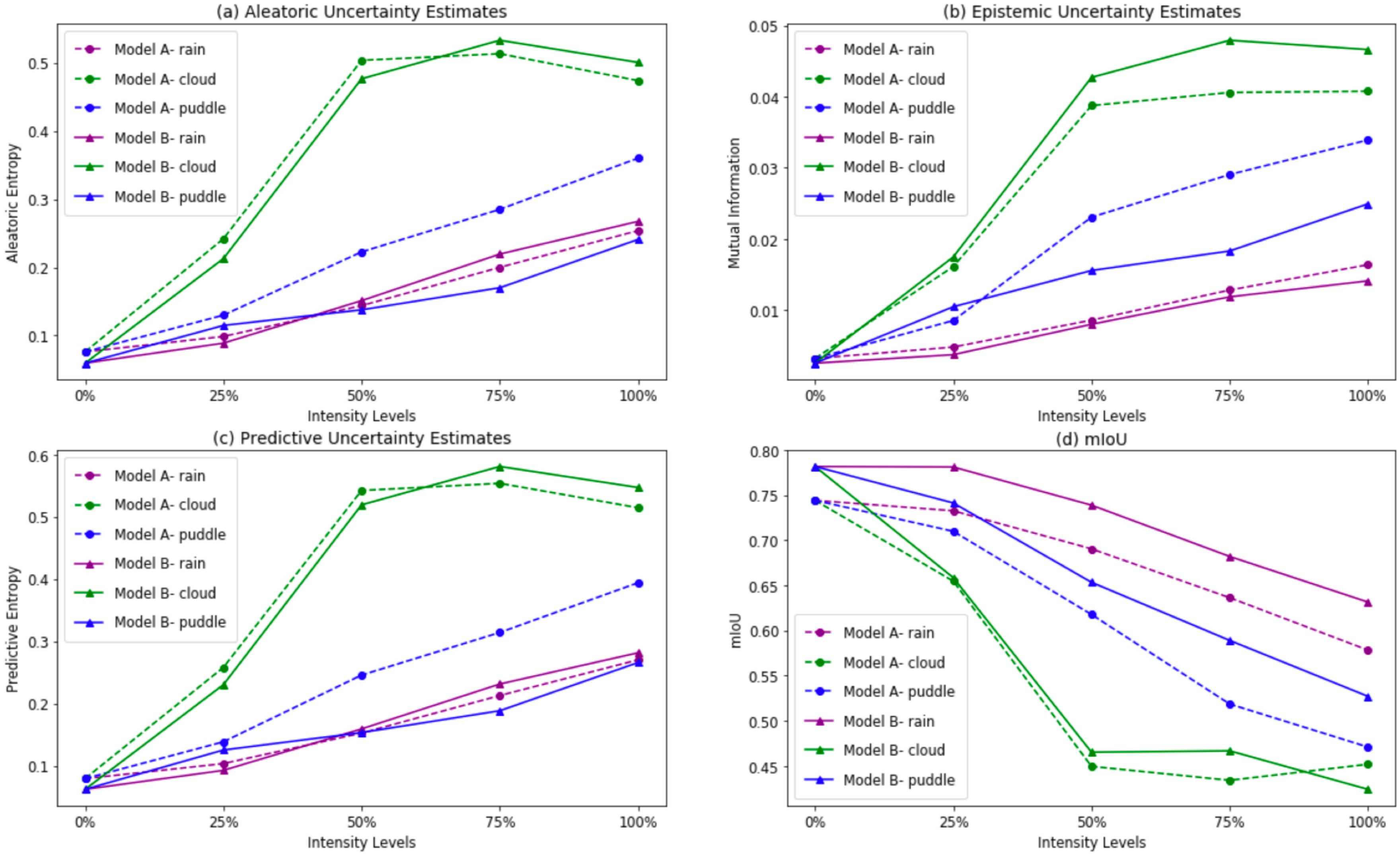
Uncertainty Estimates on Synthetic Dataset



Occlusion and Depth -> Uncertainty Estimates

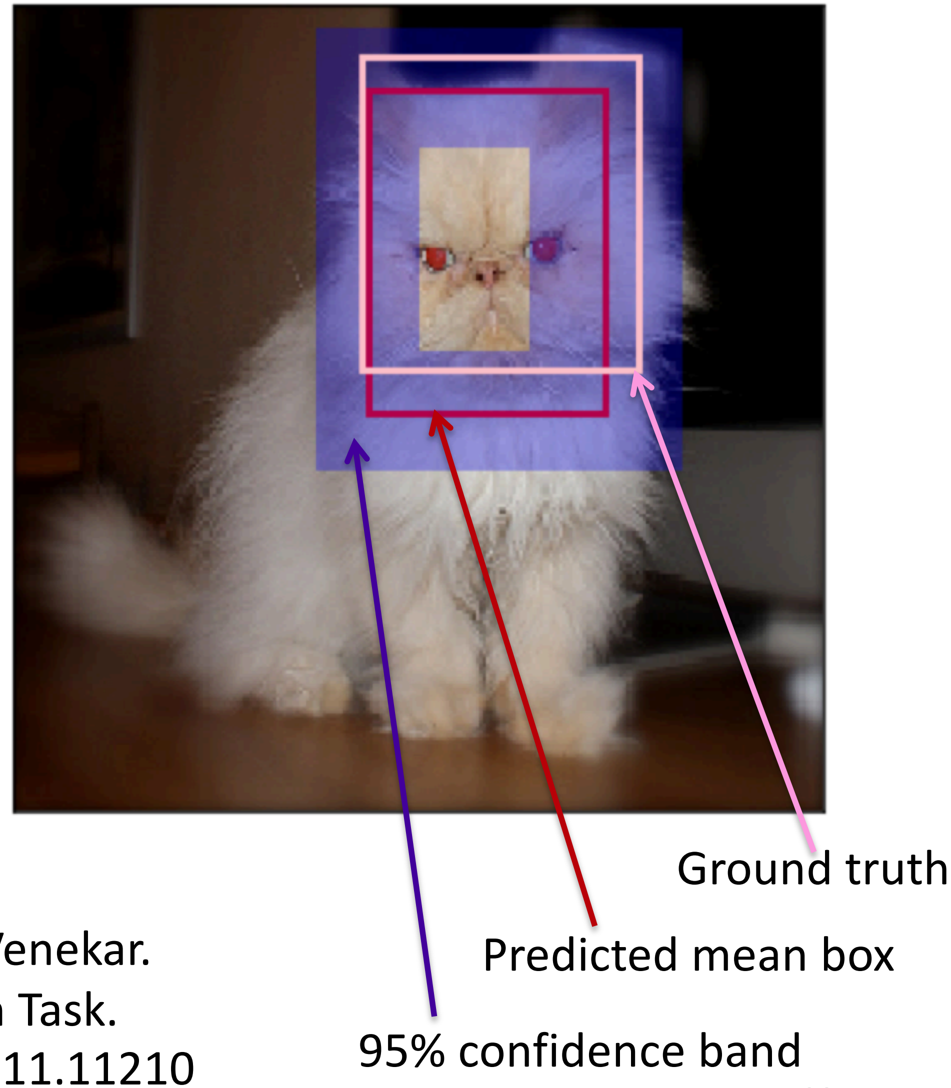


Rain, Clouds, Puddles -> Uncertainty Estimates



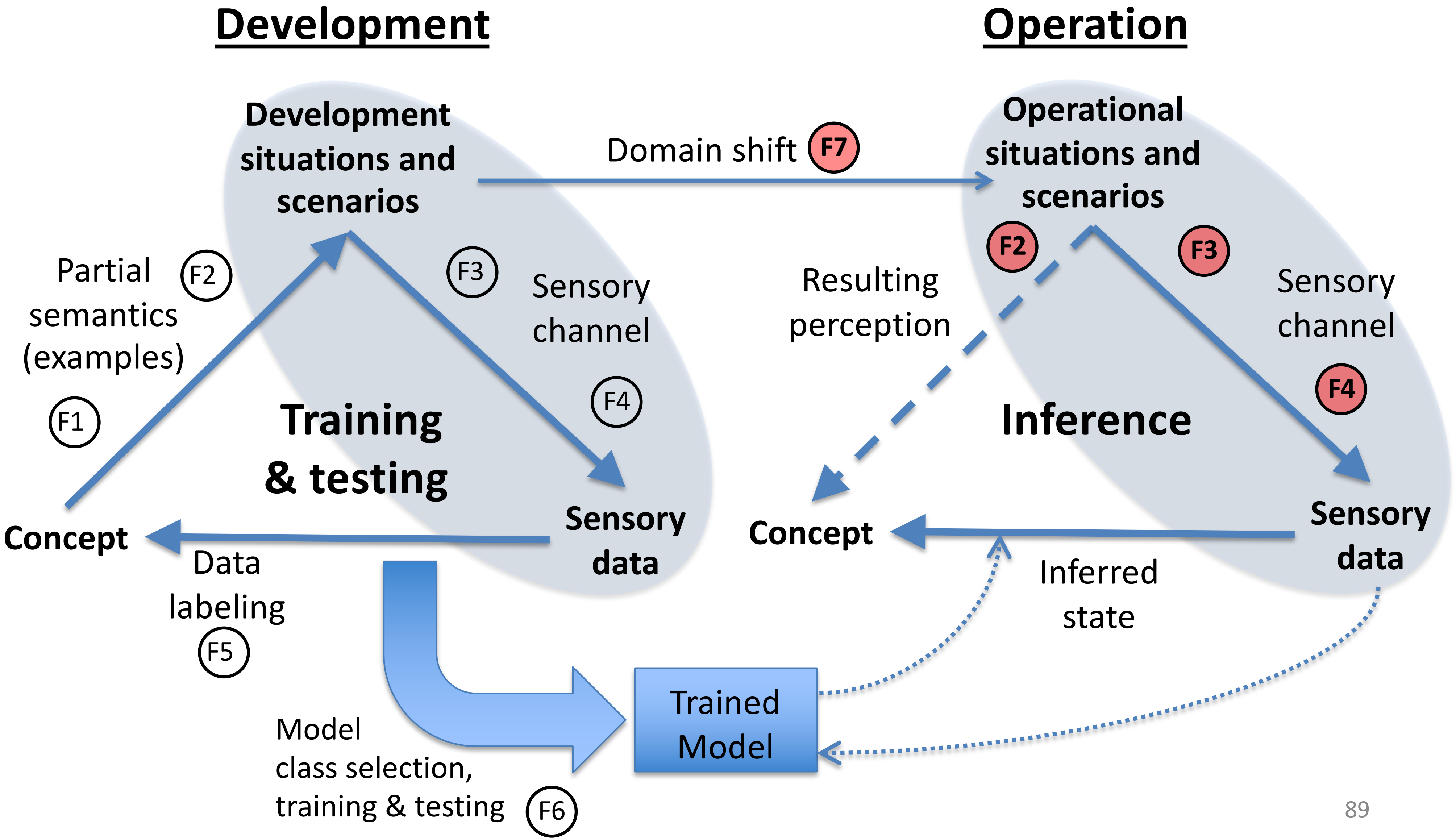
Uncertainty Estimation for Object Detection

1. Model uncertainty using MC Dropout
2. Data uncertainty using heteroschedastic regression
3. Confidence calibration



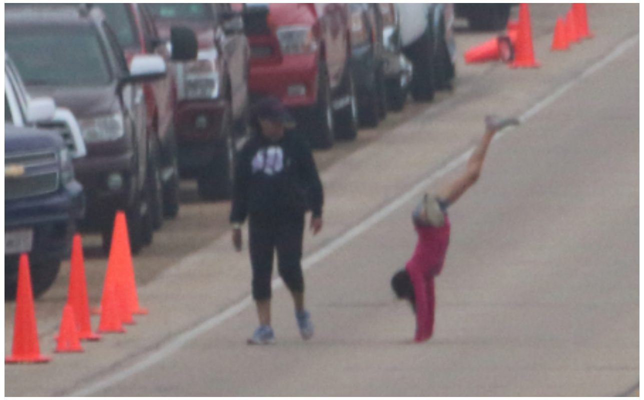
Phan, Salay, Czarnecki, Abdelzad, Denouden, Venekar. Calibrating Uncertainties in Object Localization Task. NIPS workshop. 2018, https://arxiv.org/abs/1811.11210

F7: Operational Domain Uncertainty



F7: Operational Domain Uncertainty

F2



New pedestrian pose

F3 F4

Fly splatters on LIDAR

New type of car shape

F4

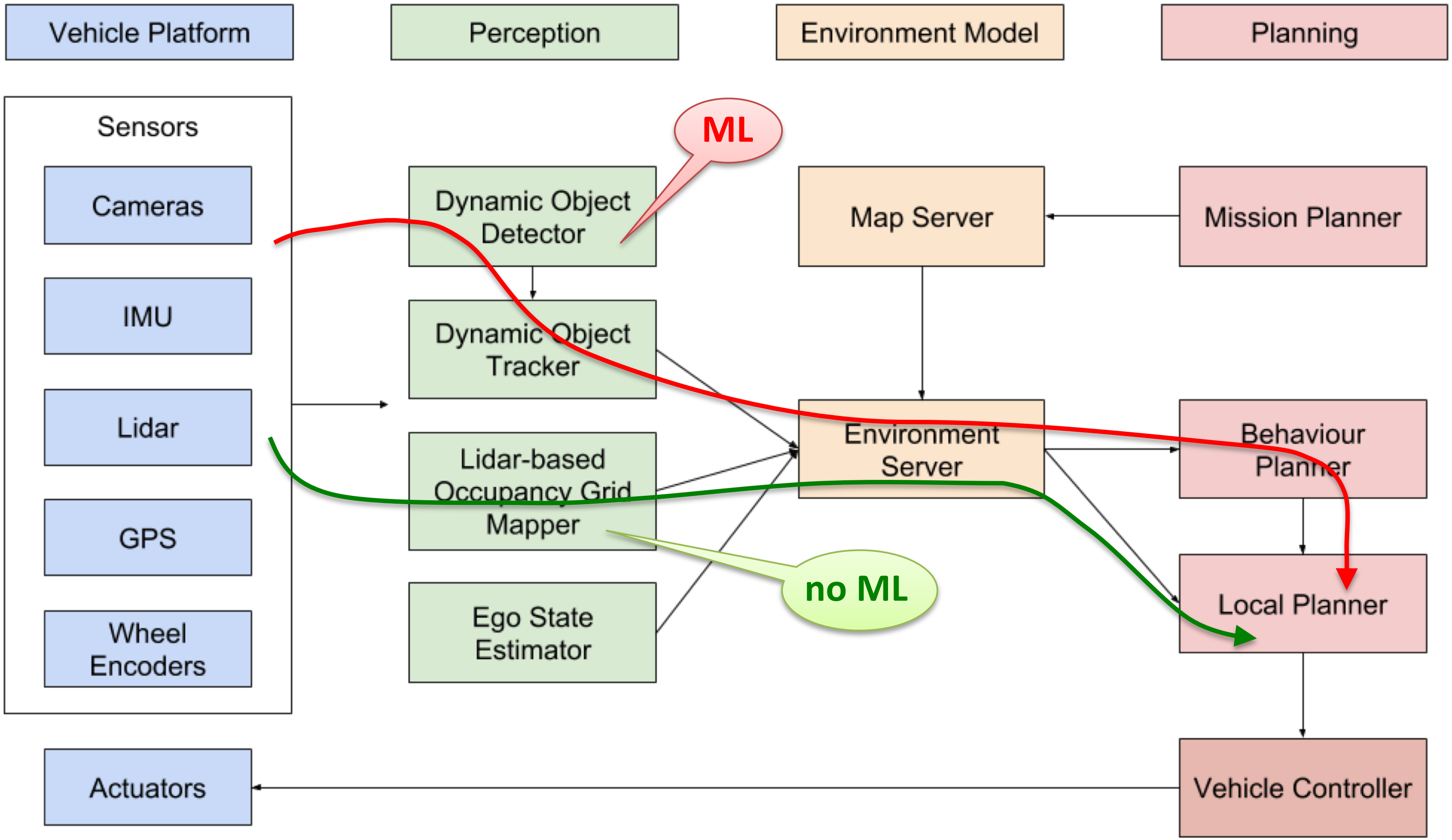
Camera miscalibration

F7: Operational Domain Uncertainty

- Assess situation novelty at operation time
  - E.g., autoencoders, partial specs
- Assess impact of level of sensor miscalibration on perceptual uncertainty
- Monitor sensor parameters and ODD

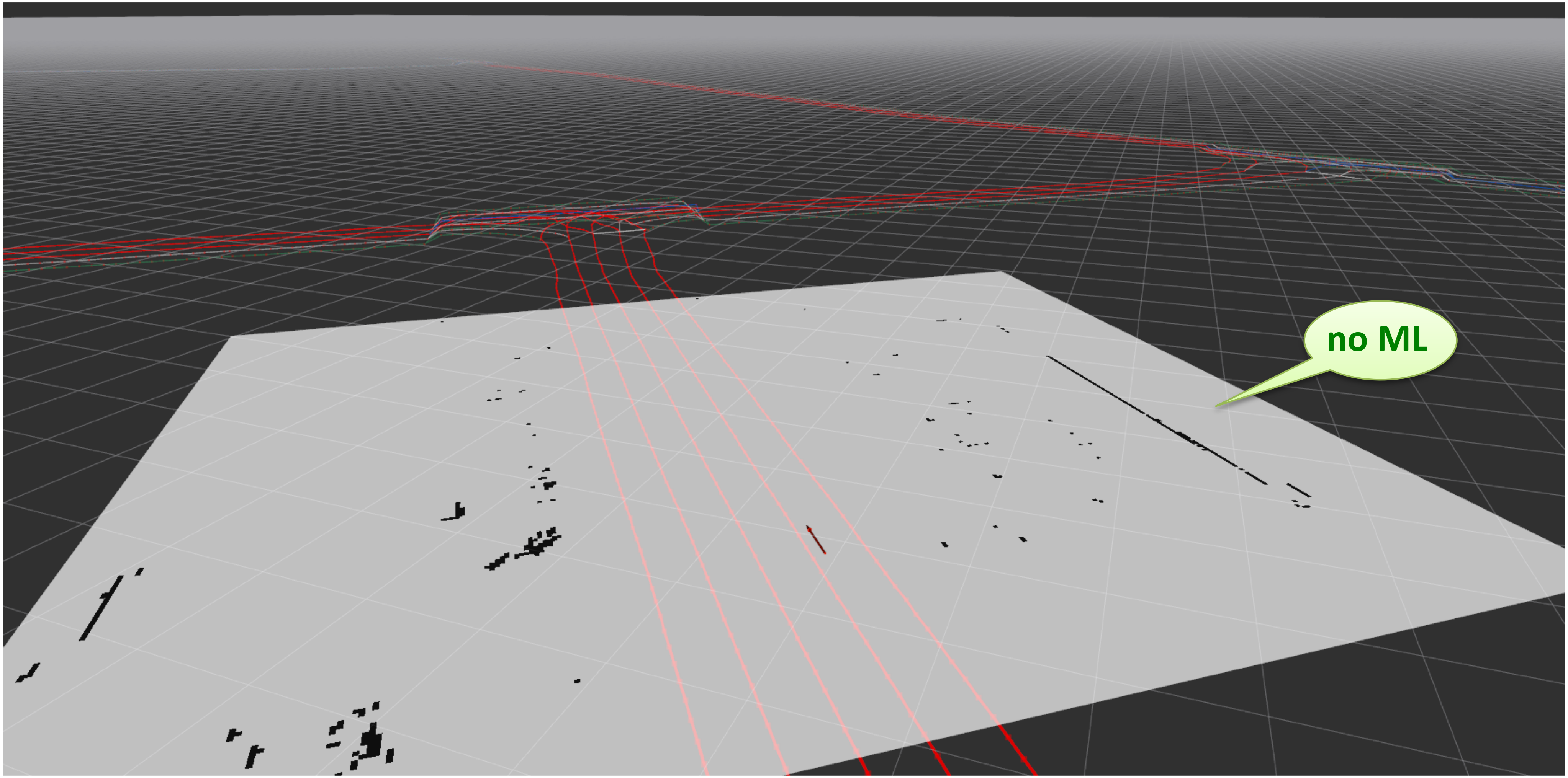
Sample Incorrect Detections

Autonomoose Architecture



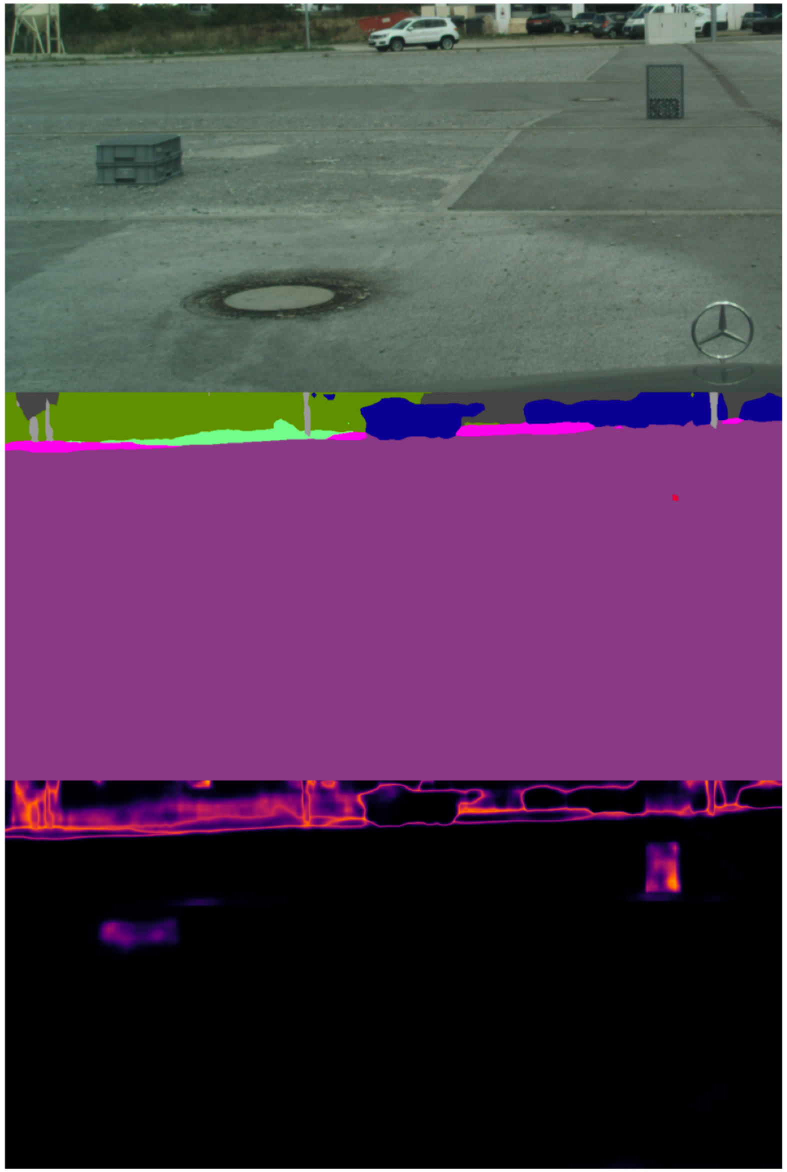
Secondary path with no ML

Lidar Occupancy Grid – Static Obstacle Detection

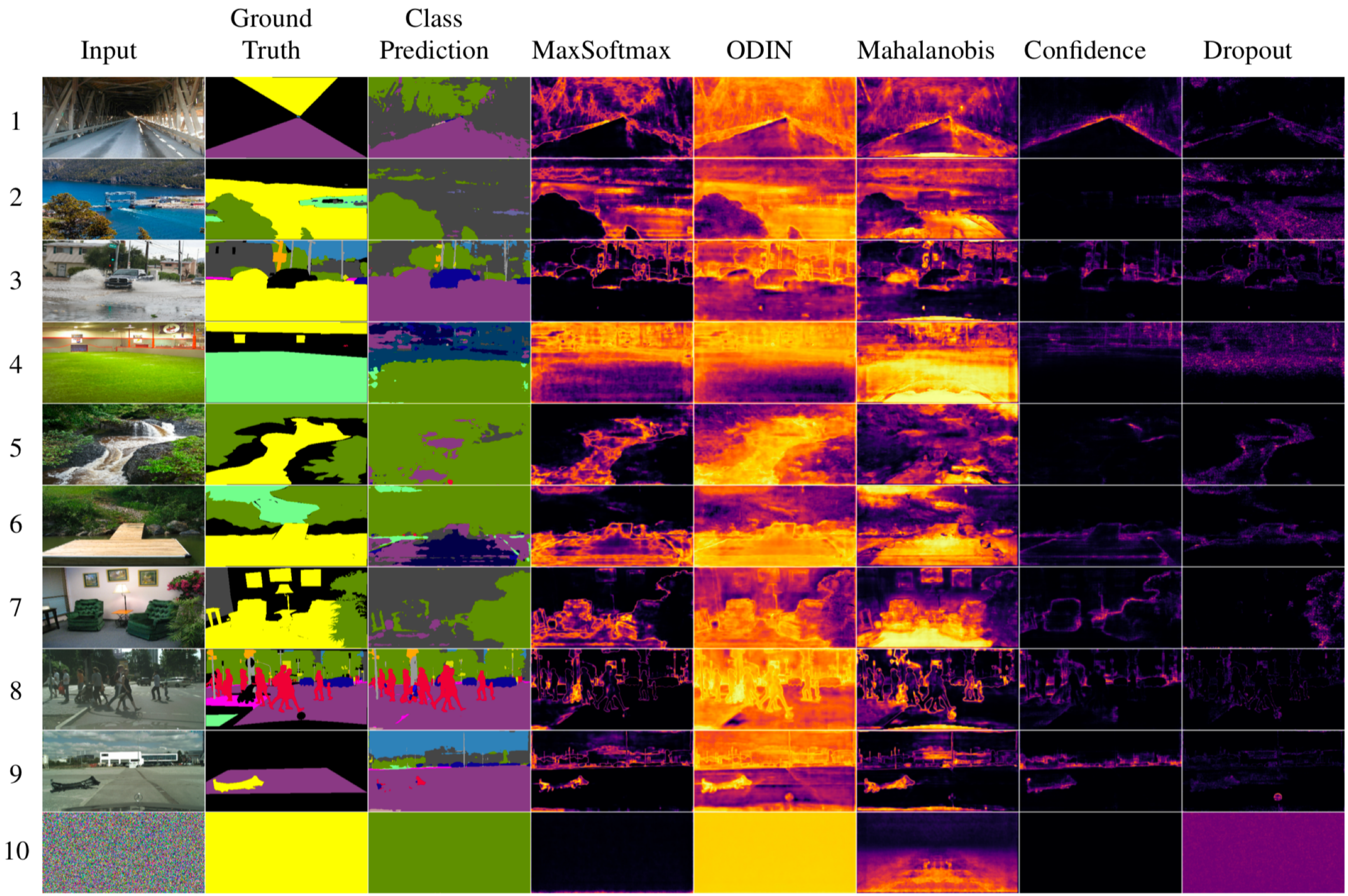


“Plastic Bag” Problem

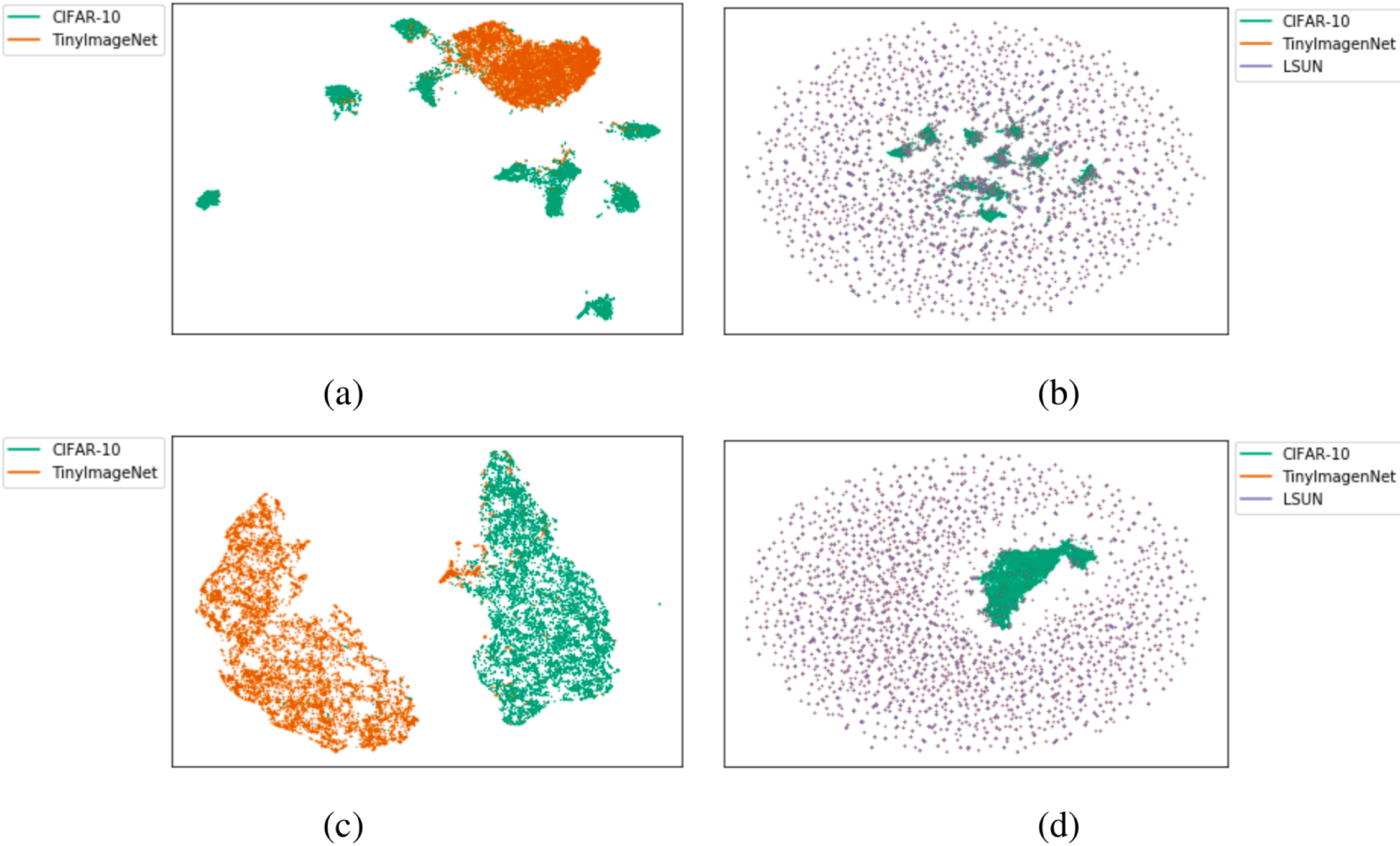
Out-Of-Distribution (OOD) for Semantic Segmentation



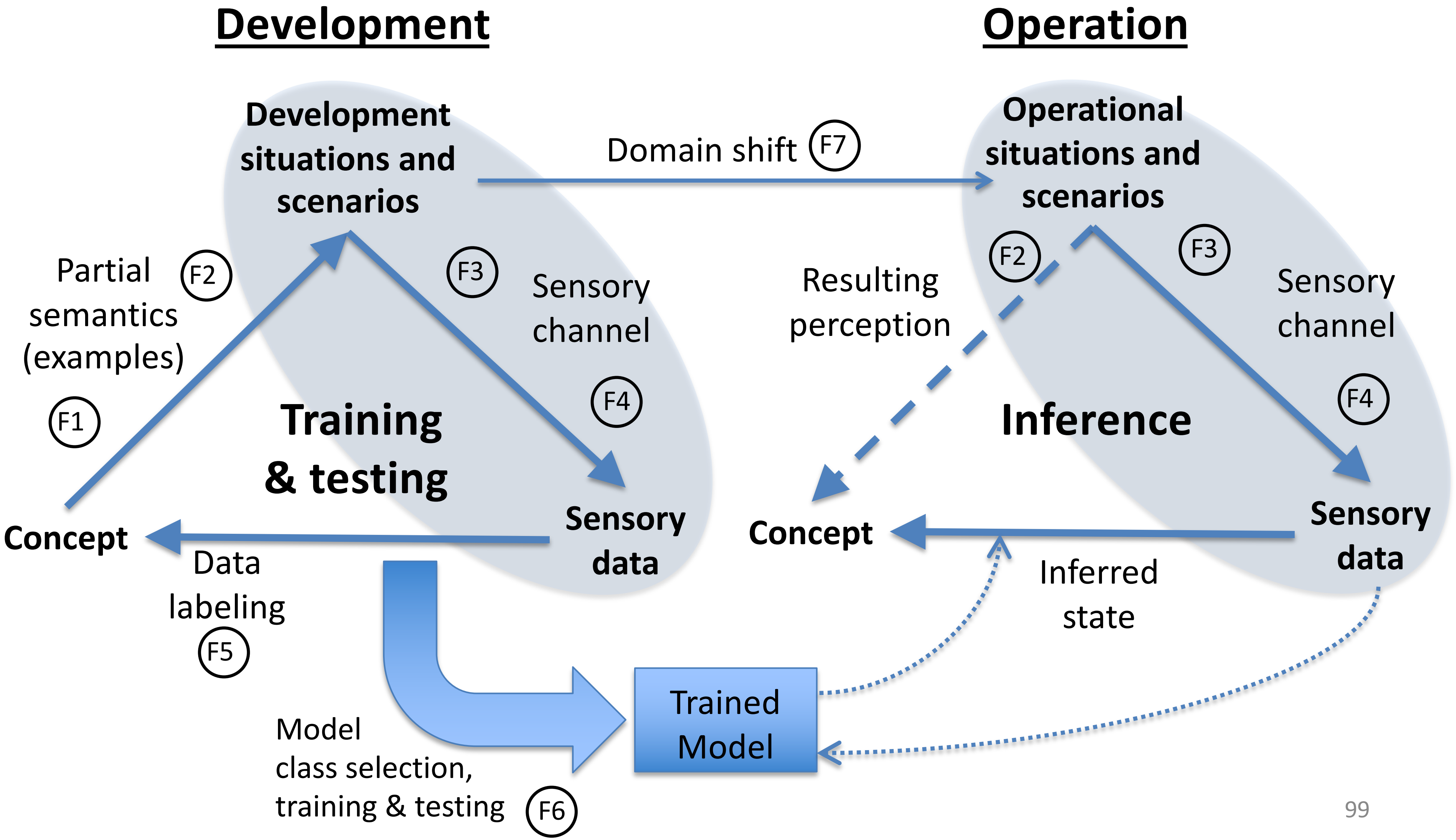
Evaluation of Five OOD Methods



New OOD Method



Factors Influencing Uncertainty

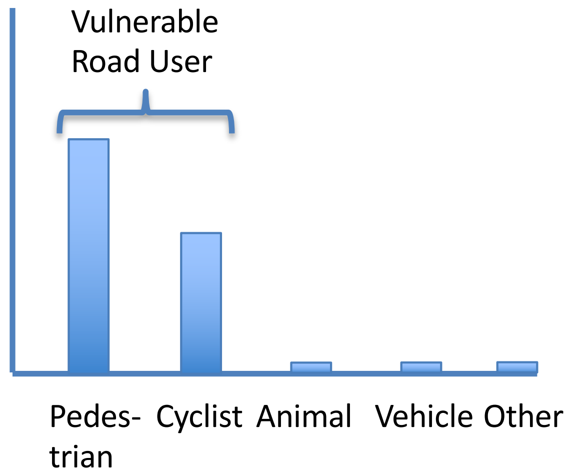


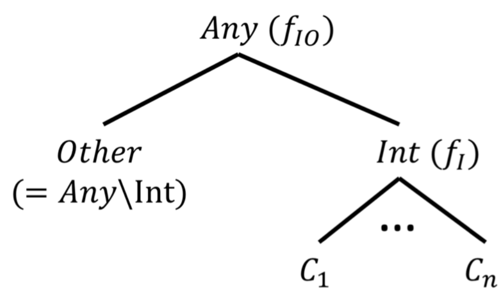
Hazard Analysis and Risk Assessment of Perceptual Failures

- Need Failure Mode Effects Analysis (FMEA) for perceptual components
  - Must deal with uncertainty
    - Uncertainty cannot be eliminated
  - Must systematically identify all failure modes
    - Perceptual equivalent of HAZOP
  - Must assess the effects
    - Incurred risk and progress cost
- Idea: introduce P-FMEA – a family of FMEAs for different perception tasks
  - C-FMEA for classification, R-FMEA for regression, OD-FMEA for object detection, etc.

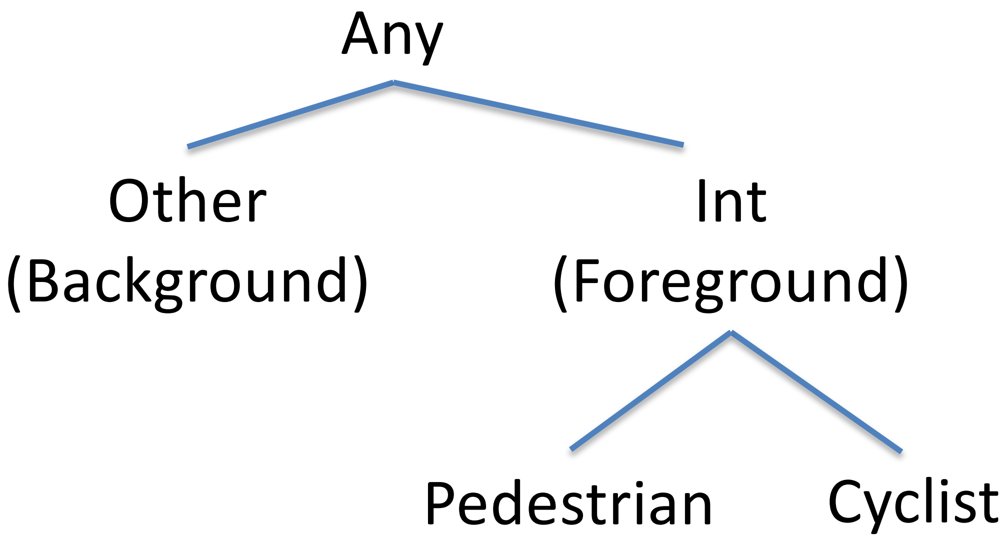
C-FMEA – Key Ideas

- Dealing with uncertainty
  - **Abstract classes** provide a more tractable representation of uncertainty than categorical distributions
- Systematic failure mode identification
  - Confusion matrix
  - **Classification case** taxonomy
- Effect analysis
  - Incurred risk and progress cost wrt. **driving policy**

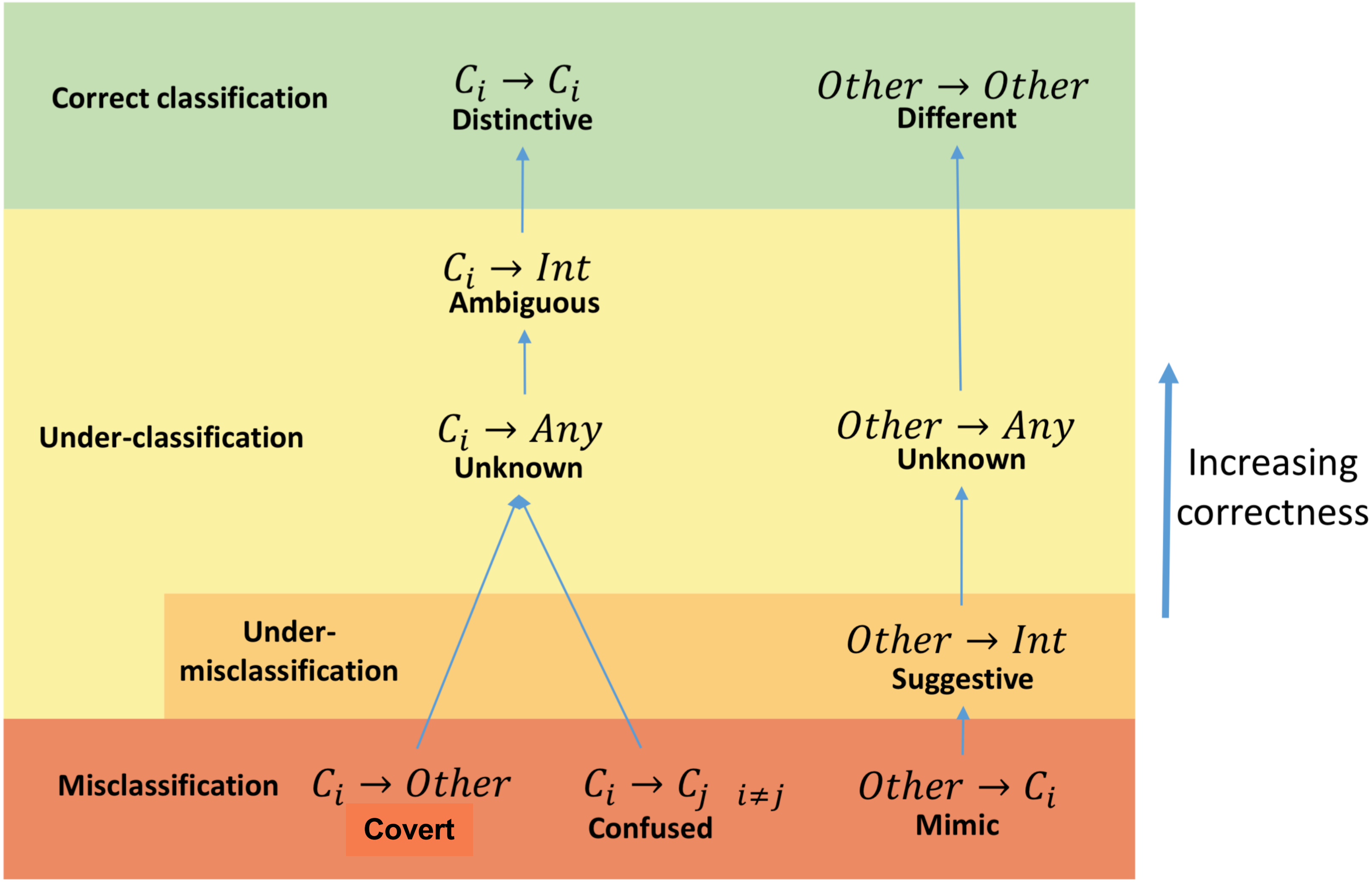




Sample Class Hierarchy



Classification Cases
(Safety-Related Order)



Classification Cases
(Safety-Related Order)

Correct classification

$C_i \rightarrow C_i$
Distinctive

$C_i \rightarrow Int$
Ambiguous

Under-classification

$C_i \rightarrow Any$
Unknown

Under-misclassification

$Other \rightarrow Int$
Suggestive

Misclassification

$C_i \rightarrow Other$
Covert

$C_i \rightarrow C_j \quad i \neq j$
Confused

$Other \rightarrow C_i$
Mimic

Pedestrian -> Pedestrian

Classification Cases
(Safety-Related Order)

Correct classification — $C_i \to C_i$ **Distinctive** — Oth

Under-classification

$C_i \to Int$ **Ambiguous**

$C_i \to Any$ **Unknown**

Under-misclassification — $Other \to Int$ **Suggestive**

Misclassification — $C_i \to Other$ **Covert** — $C_i \to C_j \quad i \neq j$ **Confused** — $Other \to C_i$ **Mimic**

Pedestrian
-> Int (Pedestrian or Cyclist)

Classification Cases
(Safety-Related Order)

Correct classification

$C_i \rightarrow C_i$
Distinctive

Oth

$C_i \rightarrow Int$
Ambiguous

Under-classification

$C_i \rightarrow Any$
Unknown

Under-misclassification

O

Misclassification

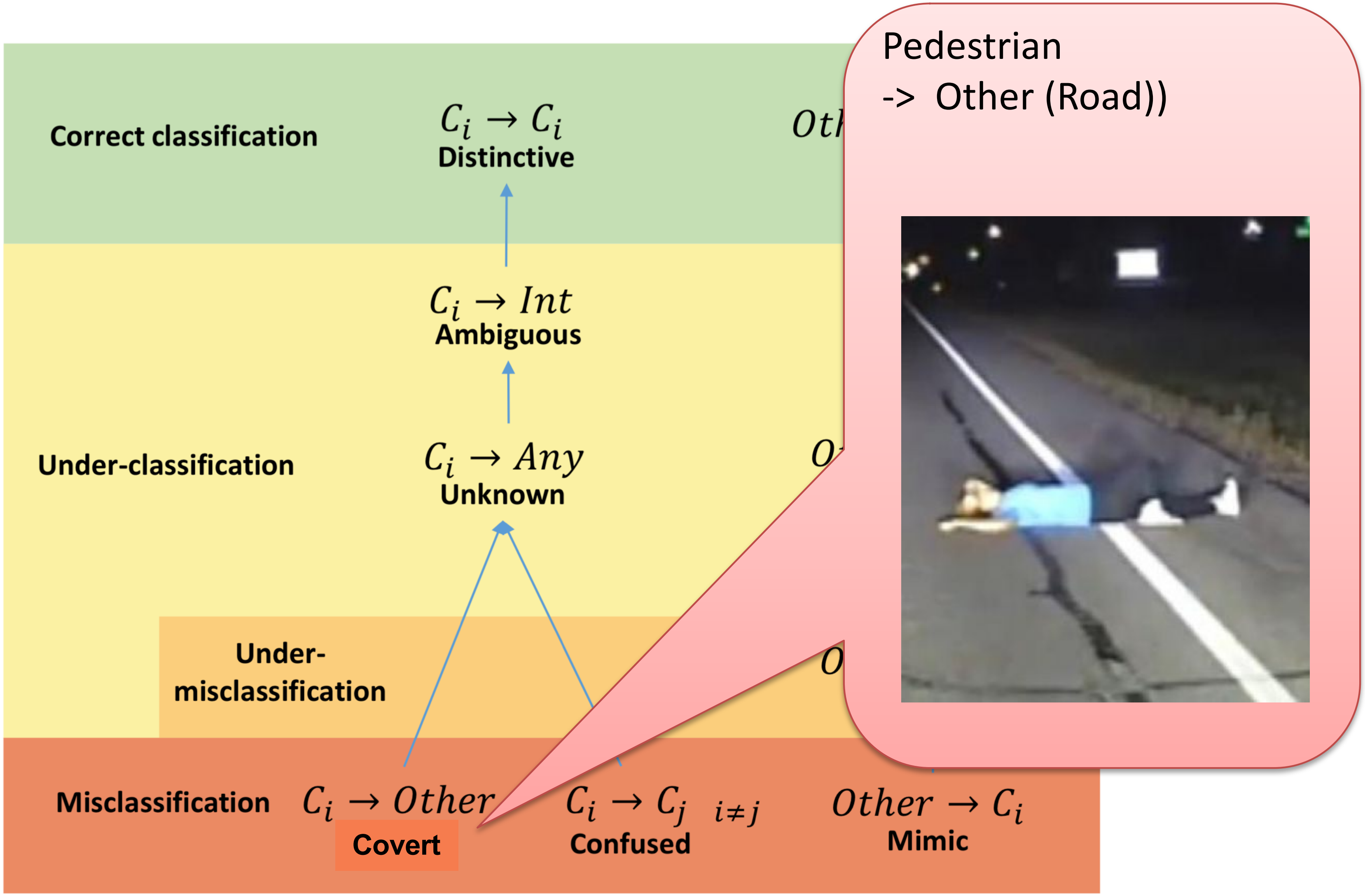
$C_i \rightarrow Other$
Covert

$C_i \rightarrow C_j \quad i \neq j$
Confused

$Other \rightarrow C_i$
Mimic

Pedestrian
-> Any (Pedestrian or
Other (Road))

Classification Cases
(Safety-Related Order)



Classification Cases
(Safety-Related Order)

Correct classification

$C_i \rightarrow C_i$
Distinctive

$C_i \rightarrow Int$
Ambiguous

Under-classification

$C_i \rightarrow Any$
Unknown

Under-misclassification

Misclassification

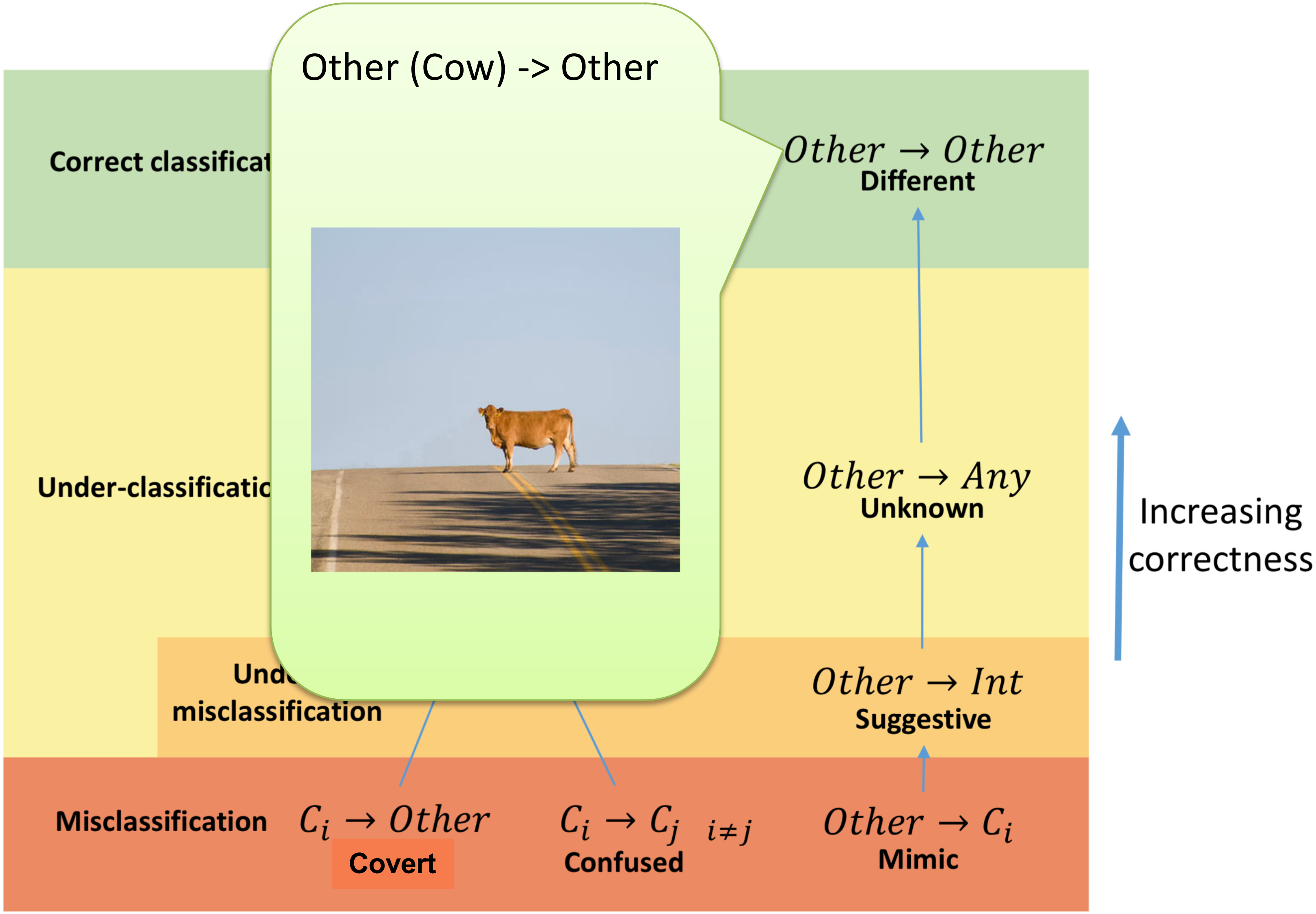
$C_i \rightarrow Other$
Covert

$C_i \rightarrow C_j \quad i \neq j$
Confused

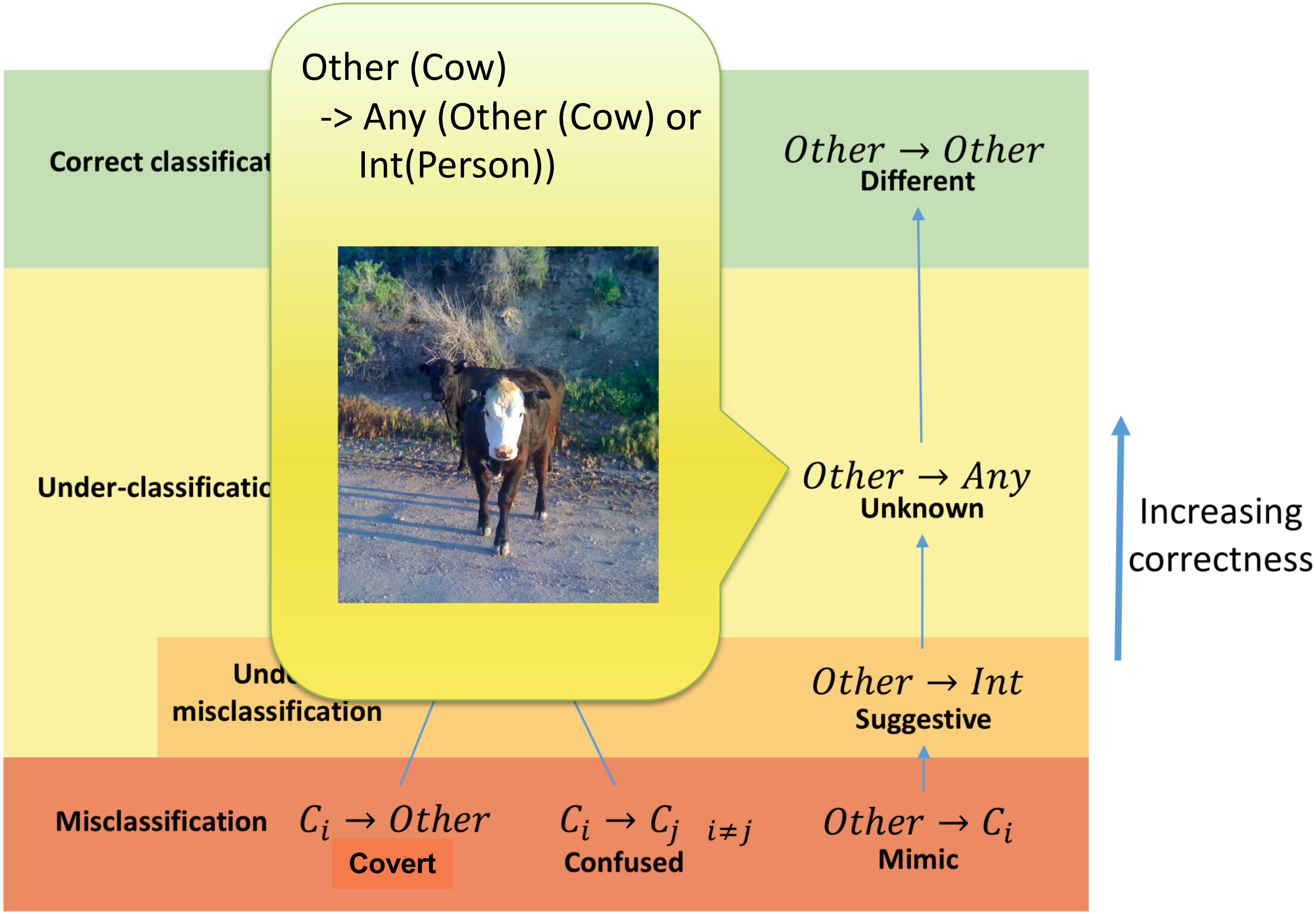
$Other \rightarrow C_i$
Mimic

Pedestrian -> Cyclist

Classification Cases
(Safety-Related Order)



Classification Cases
(Safety-Related Order)



Classification Cases
(Safety-Related Order)

Other (Statue)
-> Int (Pedestrian or Cyclist)

Correct classificat[ion] | $Other \rightarrow Other$ Different

Under-classificati[on] | $Other \rightarrow Any$ Unknown

Und[er-] misclassification | $Other \rightarrow Int$ Suggestive

Misclassification | $C_i \rightarrow Other$ Covert | $C_i \rightarrow C_j \quad i \neq j$ Confused | $Other \rightarrow C_i$ Mimic

Increasing correctness

Classification Cases
(Safety-Related Order)

Other (Manequin)
-> Pedestrian

Correct classificat

$Other \rightarrow Other$
Different

Under-classificatio

$Other \rightarrow Any$
Unknown

Und
misclassification

$Other \rightarrow Int$
Suggestive

Misclassification

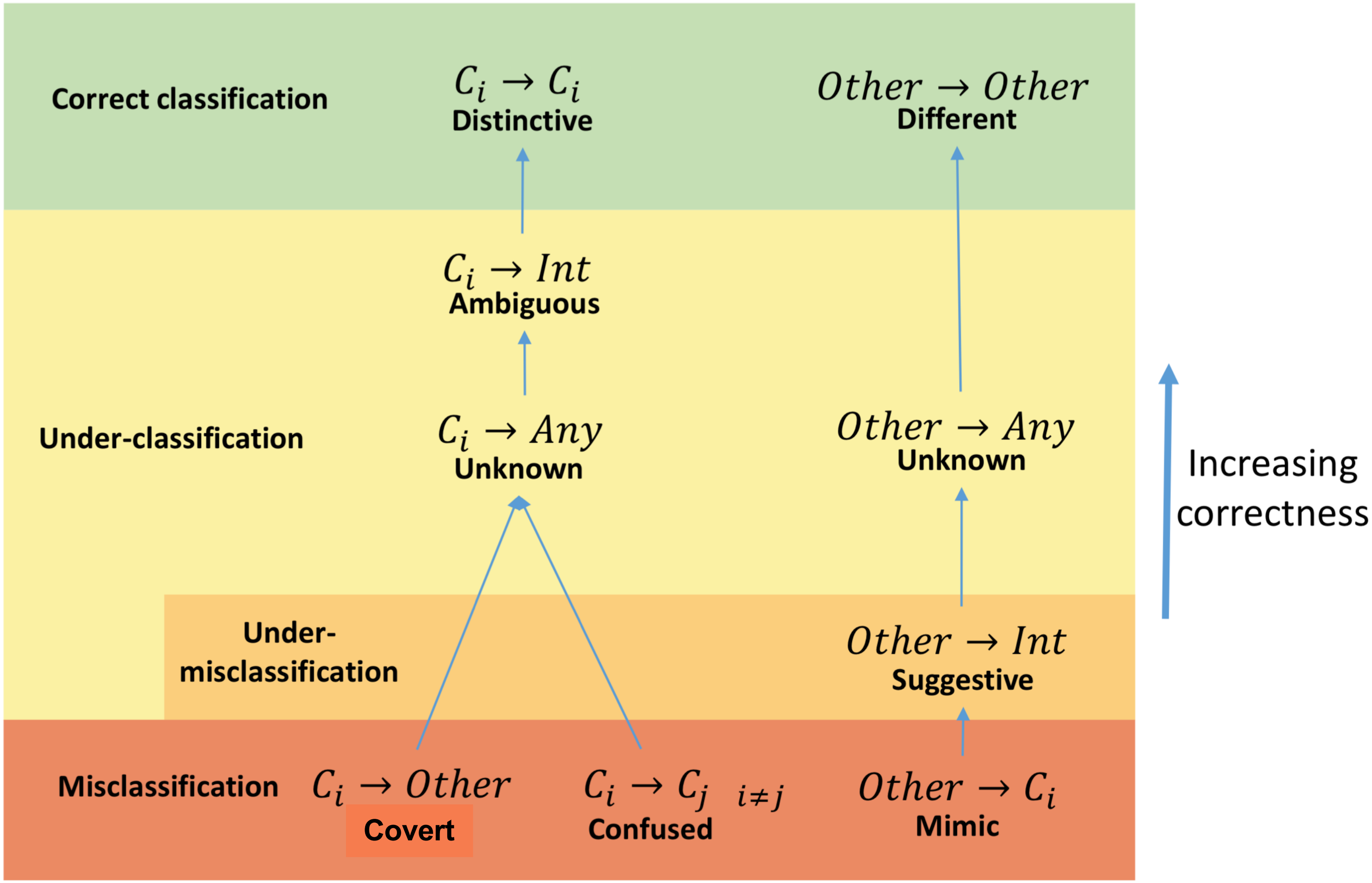
$C_i \rightarrow Other$
Covert

$C_i \rightarrow C_j \quad i \neq j$
Confused

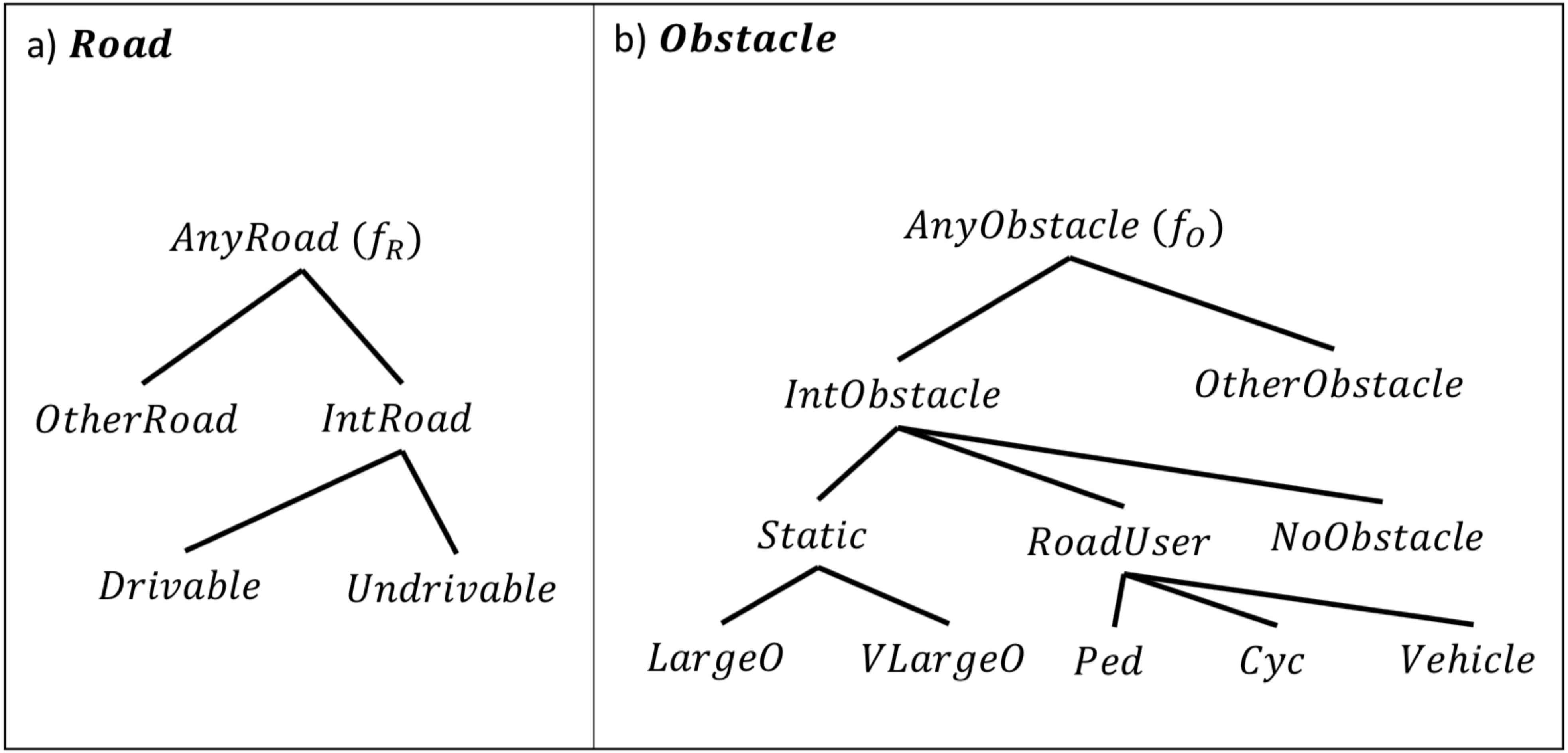
$Other \rightarrow C_i$
Mimic

Increasing correctness

Classification Cases
(Safety-Related Order)



Case Study – Class Hierarchy



Case Study – Perception Module

Case Study – Driving Policy

a) π_R

Class	Action
$Drivable$	$Cruise$
$Undrivable$	$Stop$
$OtherRoad$	$Stop$
$IntRoad$	$Stop$
$AnyRoad$	$Stop$

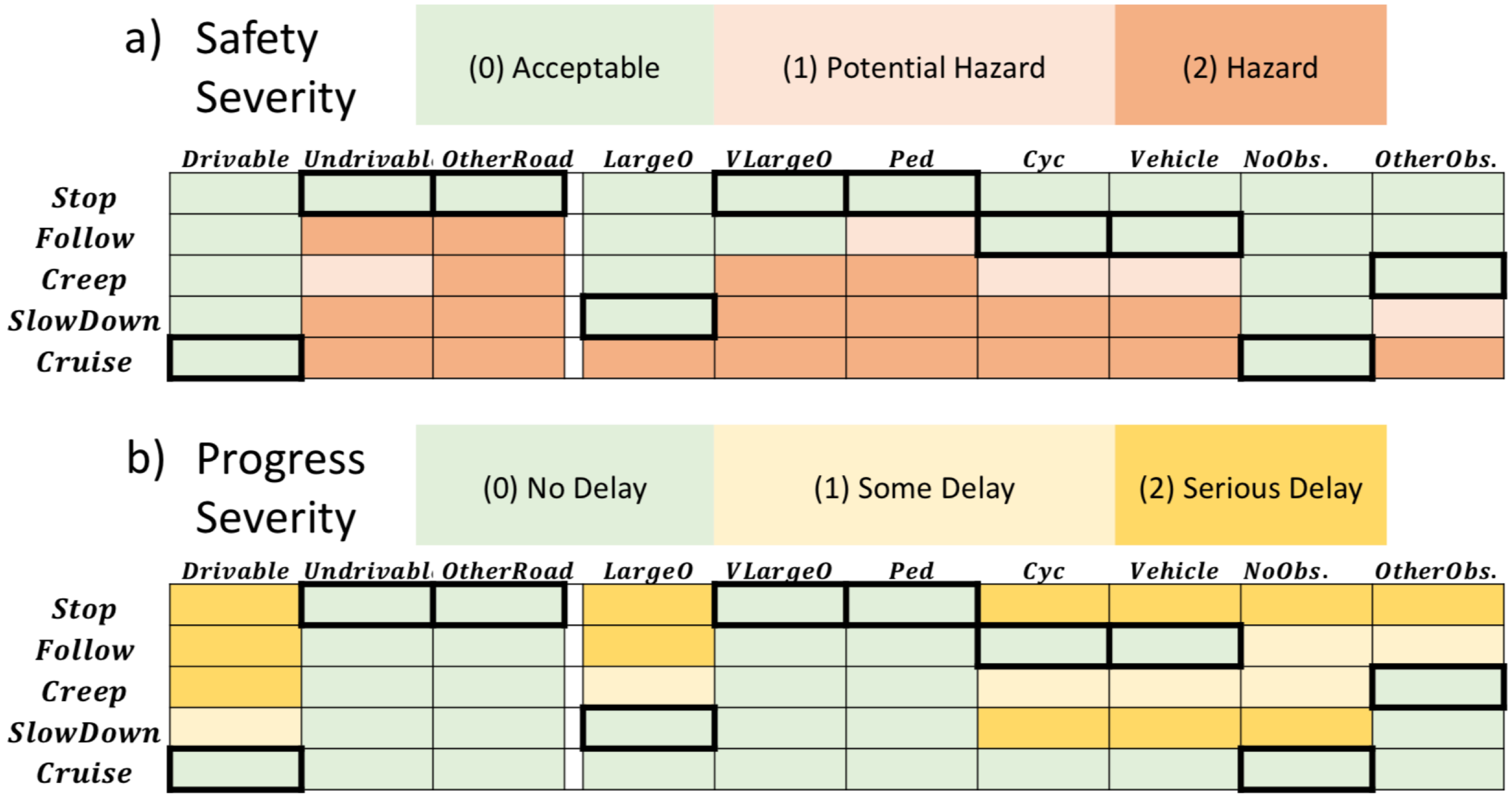
b) π_O

Class	Action
$LargeO$	$Slowdown$
$VLargeO$	$Stop$
Ped	$Stop$
Cyc	$Follow$
$Vehicle$	$Follow$
$NoObstacle$	$Cruise$
$OtherObstacle$	$Creep$
$Static$	$Stop$
$RoadUser$	$Stop$
$IntObstacle$	$Stop$
$AnyObstacle$	$Stop$

c)

$Stop$
$Follow$
$Creep$
$Slowdown$
$Cruise$

Case Study – Policy Deviation Safety and Progress Assessment

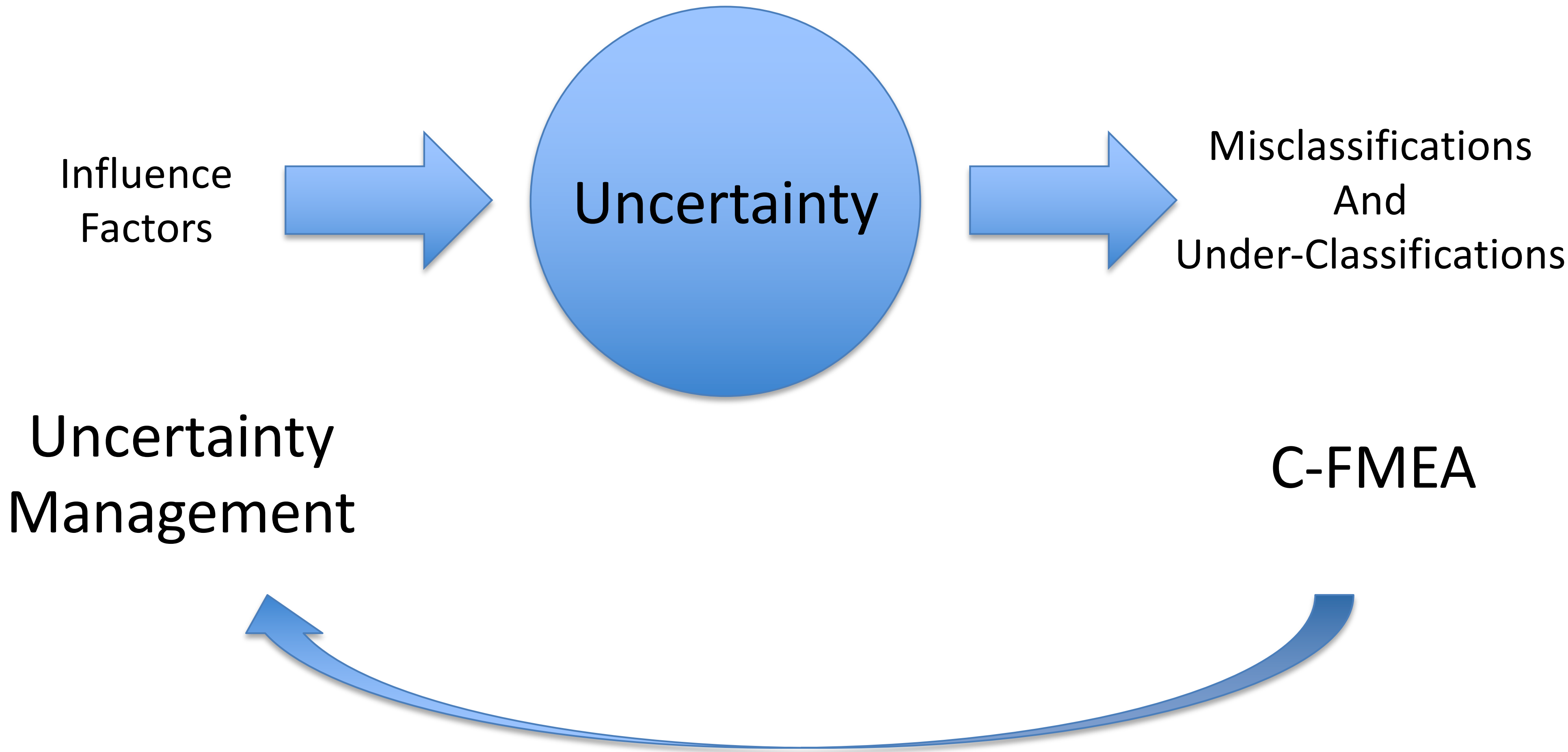


Case Study – Configuration Case Safety and Progress Assessment

	LargeO	*VLargeO*	*Ped*	*Cyc*	*Vehicle*	*NoObs.*	*OtherObs.*
LargeO	00	20	20	22	22	02	10
VLargeO	02 (0.002)	00 (0.006)	00	02	02	02 (0.004)	02
Ped	02 (0.002)	00	00 (0.048)	02	02	02 (0.004)	02
Cyc	02 (0.004)	00	10	00 (0.014)	00	01 (0.004)	01 (0.002)
Vehicle	02 (0.016)	00 (0.006)	10 (0.006)	00	00 (0.066)	01 (0.066)	01
NoObs.	20 (0.020)	20 (0.022)	20 (0.010)	20 (0.002)	20 (0.036)	00 (0.406)	20 (0.006)
OtherObs.	01	20	20	11	11	01	00
RoadUser	02 (0.004)	00	00 (0.024)	02 (0.014)	02 (0.014)	02 (0.008)	02
Static	02 (0.002)	00	00	02	02	02	02
IntObs.	02	00	00	02	02	02	02
AnyObs.	02	00	00	02	02	02	02

Confusion Matrix

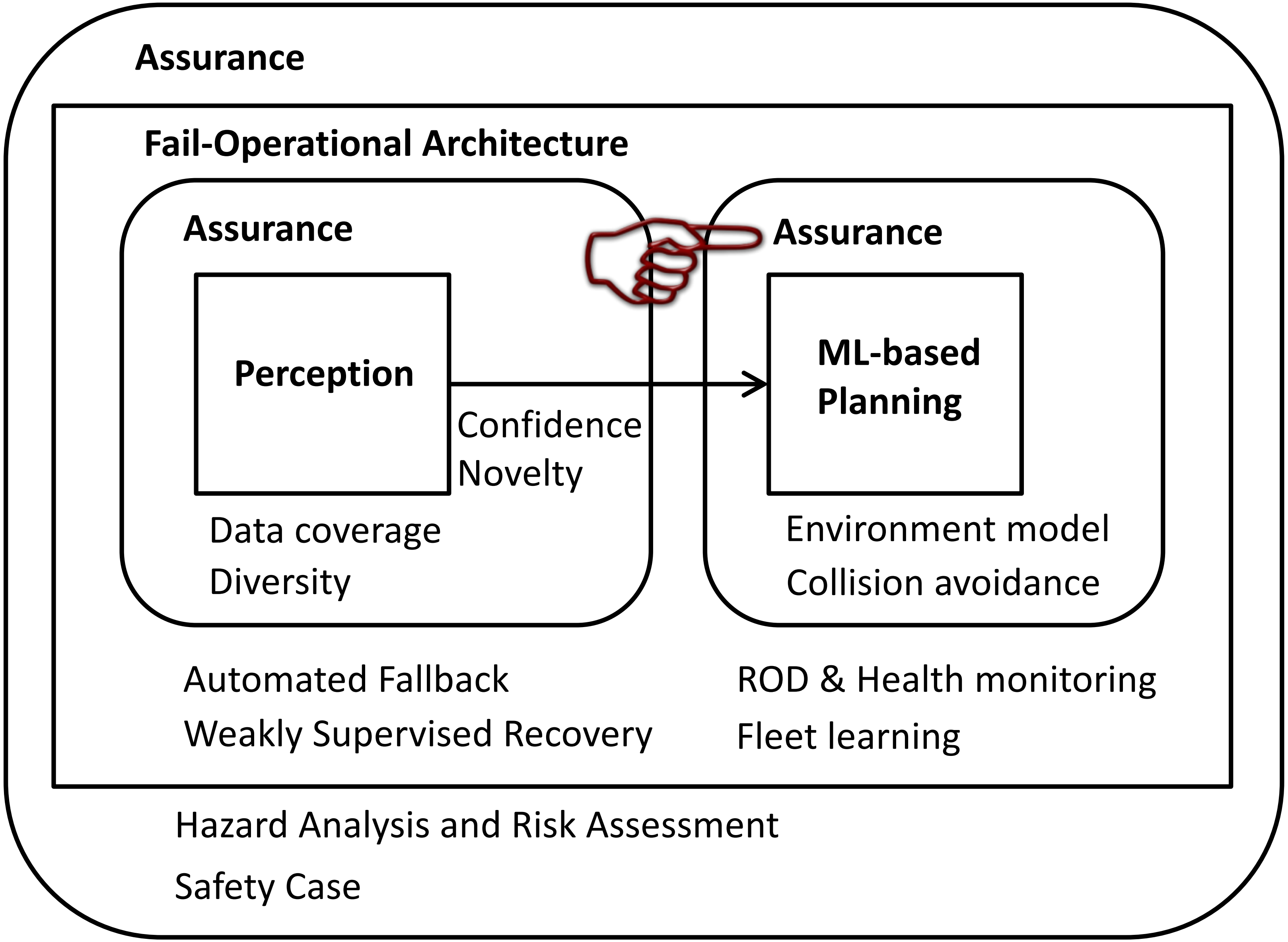
How The Ideas Fit Together?



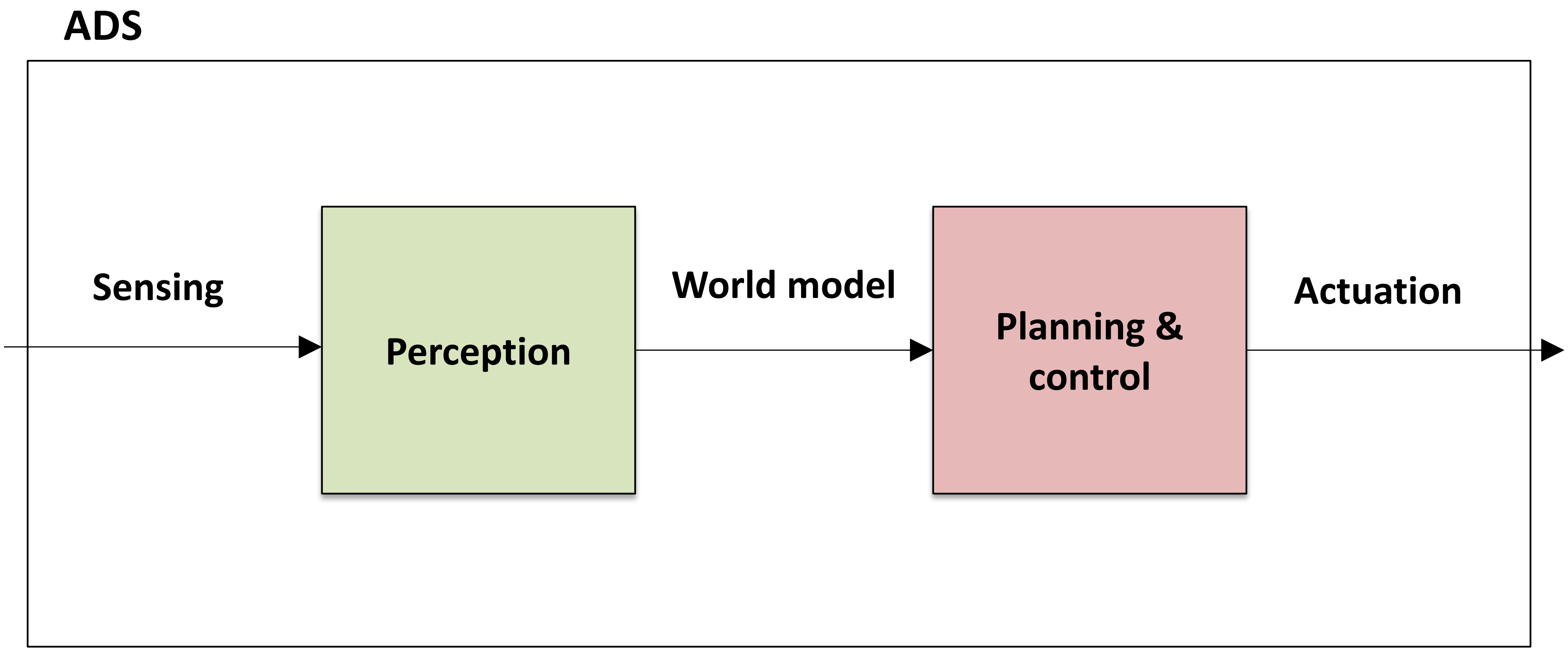
Part II Summary

1. Perceptual uncertainty is a key performance measure in safety requirements
2. Used perceptual triangle to identify seven influence factors for perceptual uncertainty when using supervised ML
3. FMEA for Perception Functions
4. Future: methods to control the influence factors and use them in safety arguments

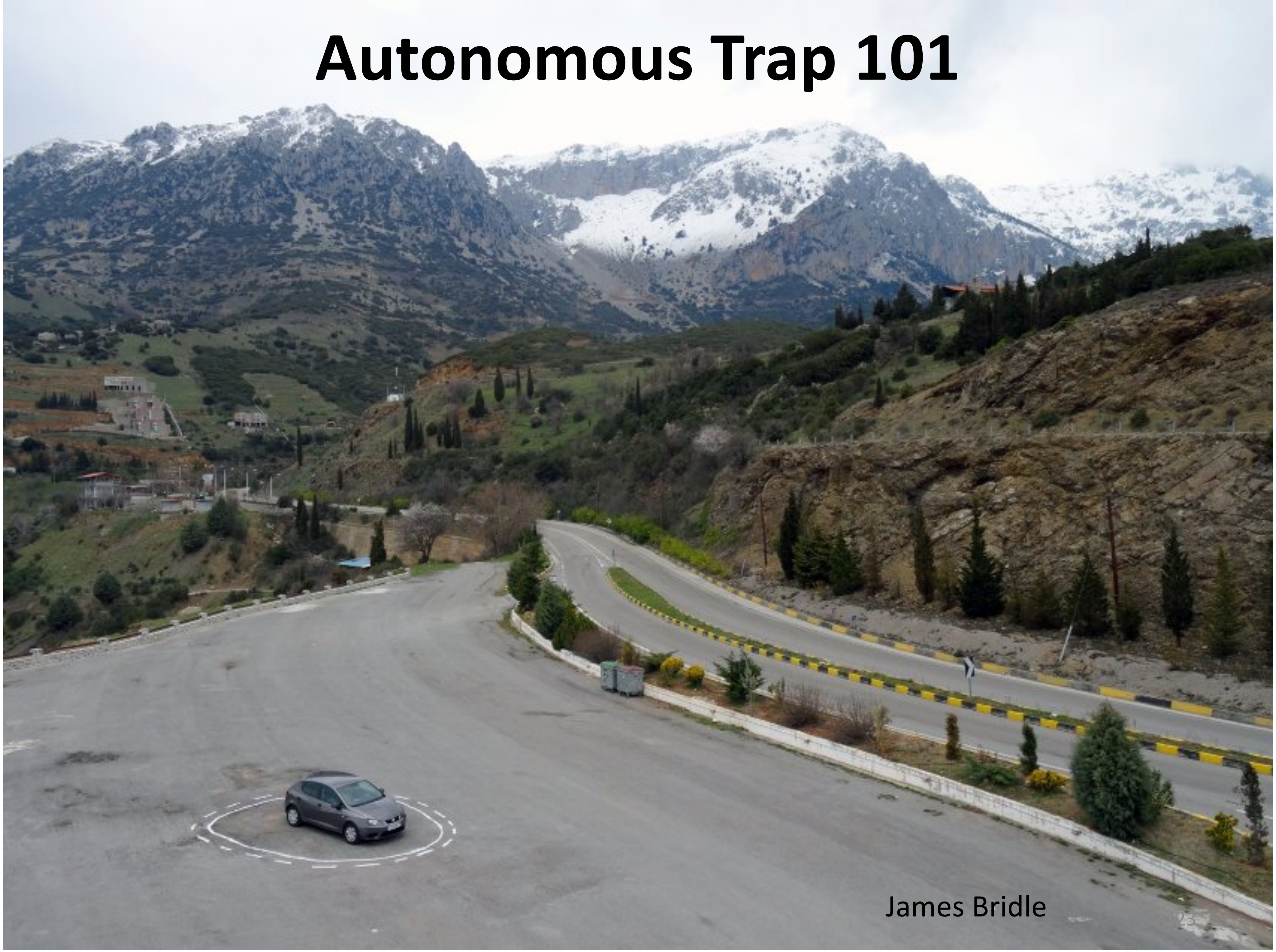
LAVA: Learned & Assured Vehicle Autonomy



Safety Argument Decomposition



Autonomous Trap 101

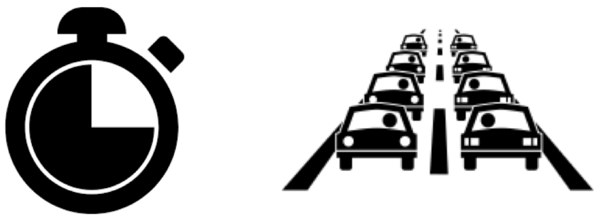


James Bridle

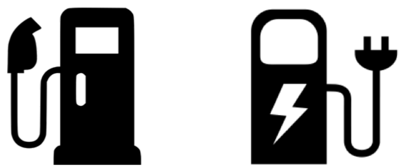
Driving Qualities

Safety

Comfort

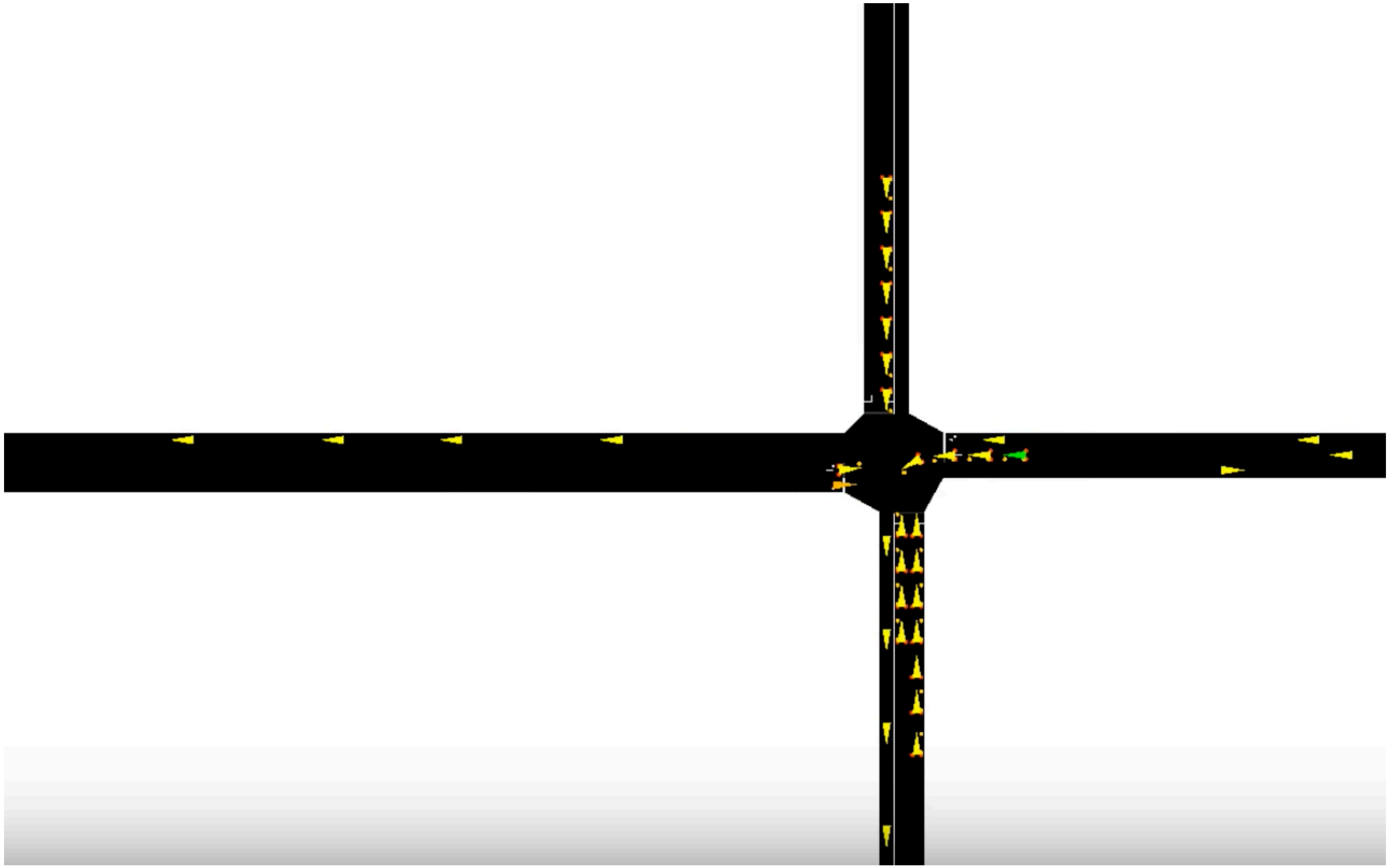


Progress



Energy efficiency

Multi-Objective Reinforcement Learning



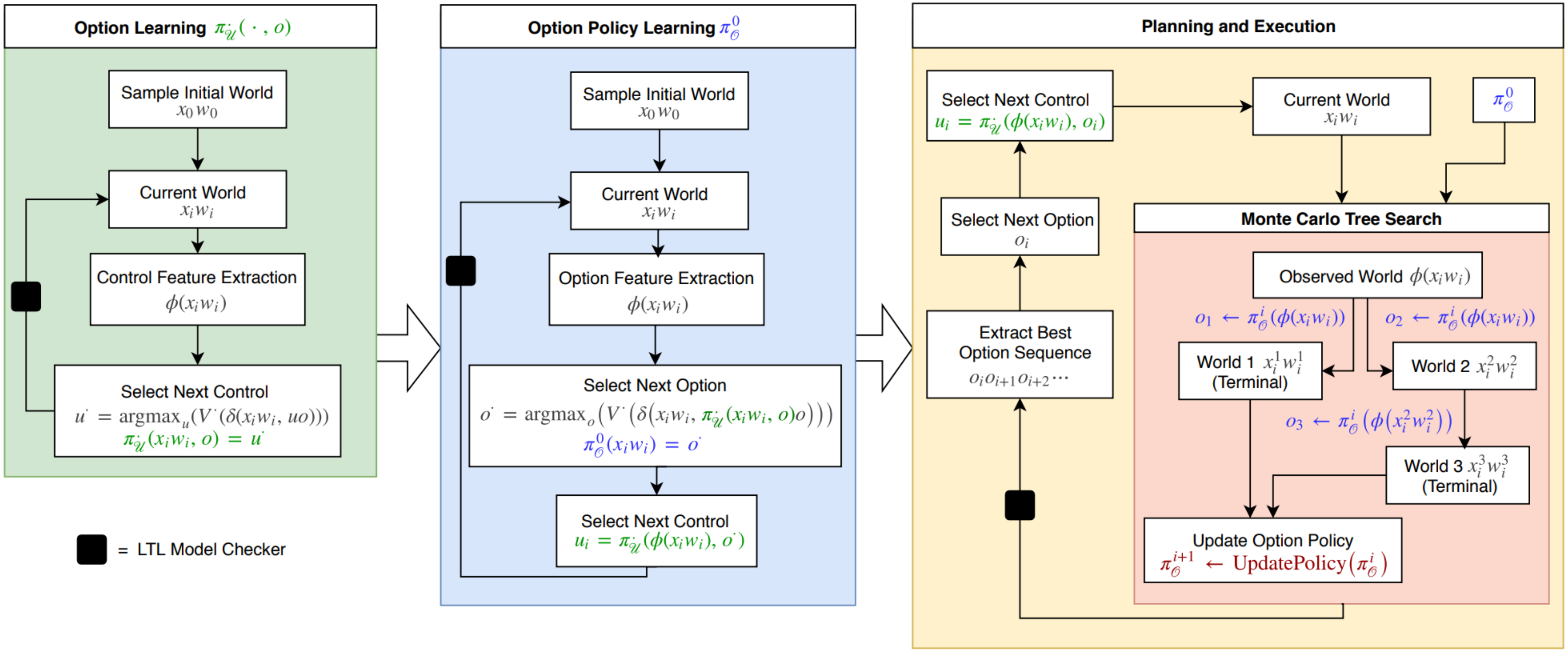
Videos: https://www.youtube.com/playlist?list=PLiZsfe-Hr4k9VPiX0tfoNoHHDUE2MDPuQ

Li et al. Urban Driving with Multi-Objective Deep Reinforcement Learning. Under review, 2018
https://arxiv.org/abs/1811.08586

Deep RL Challenges

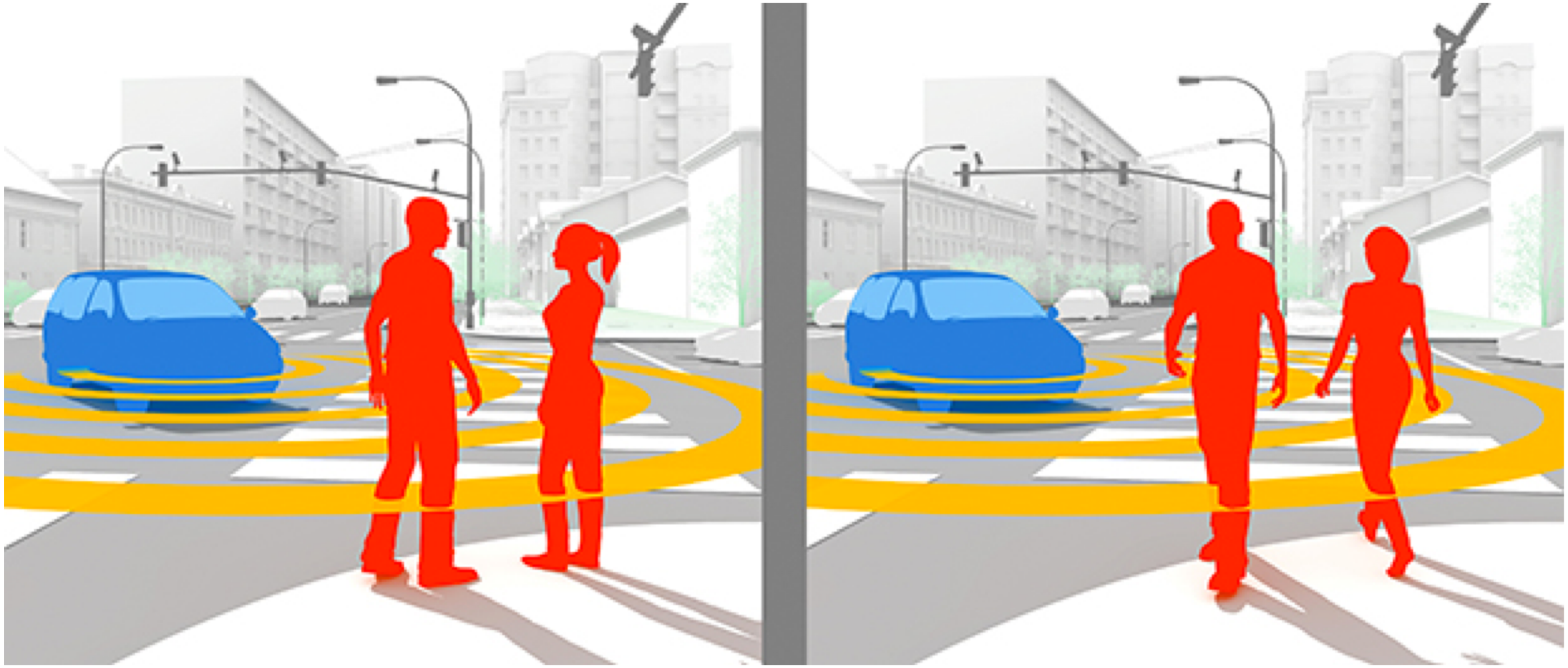
- Environment model
- Rewards and specifications
- Learning is slow
  – Should combine with imitation learning and MPC-based maneuvers
- Safety
  – Safety envelope
    - Escape path & fallback path
  – Analyzable policies

Baseline RL Architecture for Automated Driving

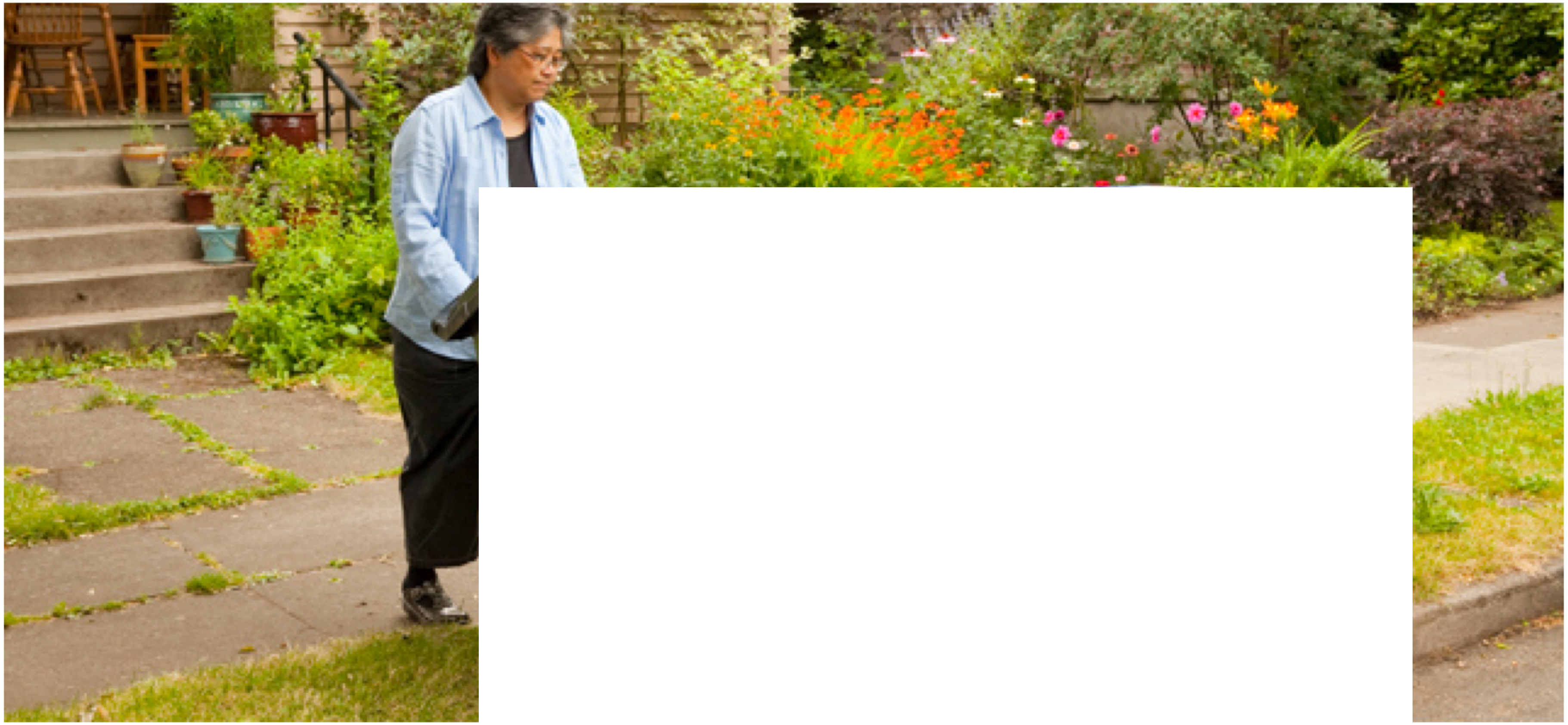


Paxton, et al. M. Combining neural networks and tree search for task and motion planning in challenging environments. arXiv preprint arXiv:1703.07887, 2017

Road User Intension



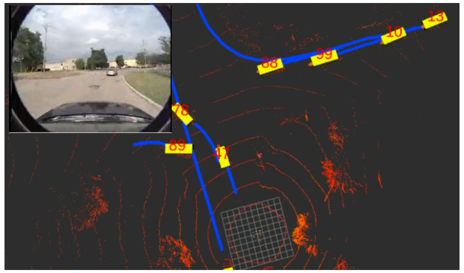
Will she cross the street?



Will she cross the street?

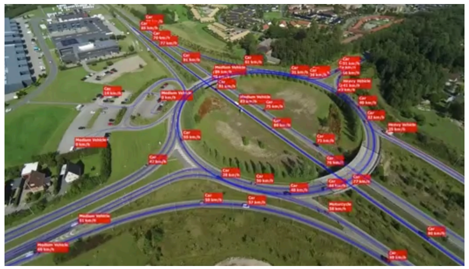
Traffic Data

Naturalistic driving



AV sensors & perception

Infrastructure mounted



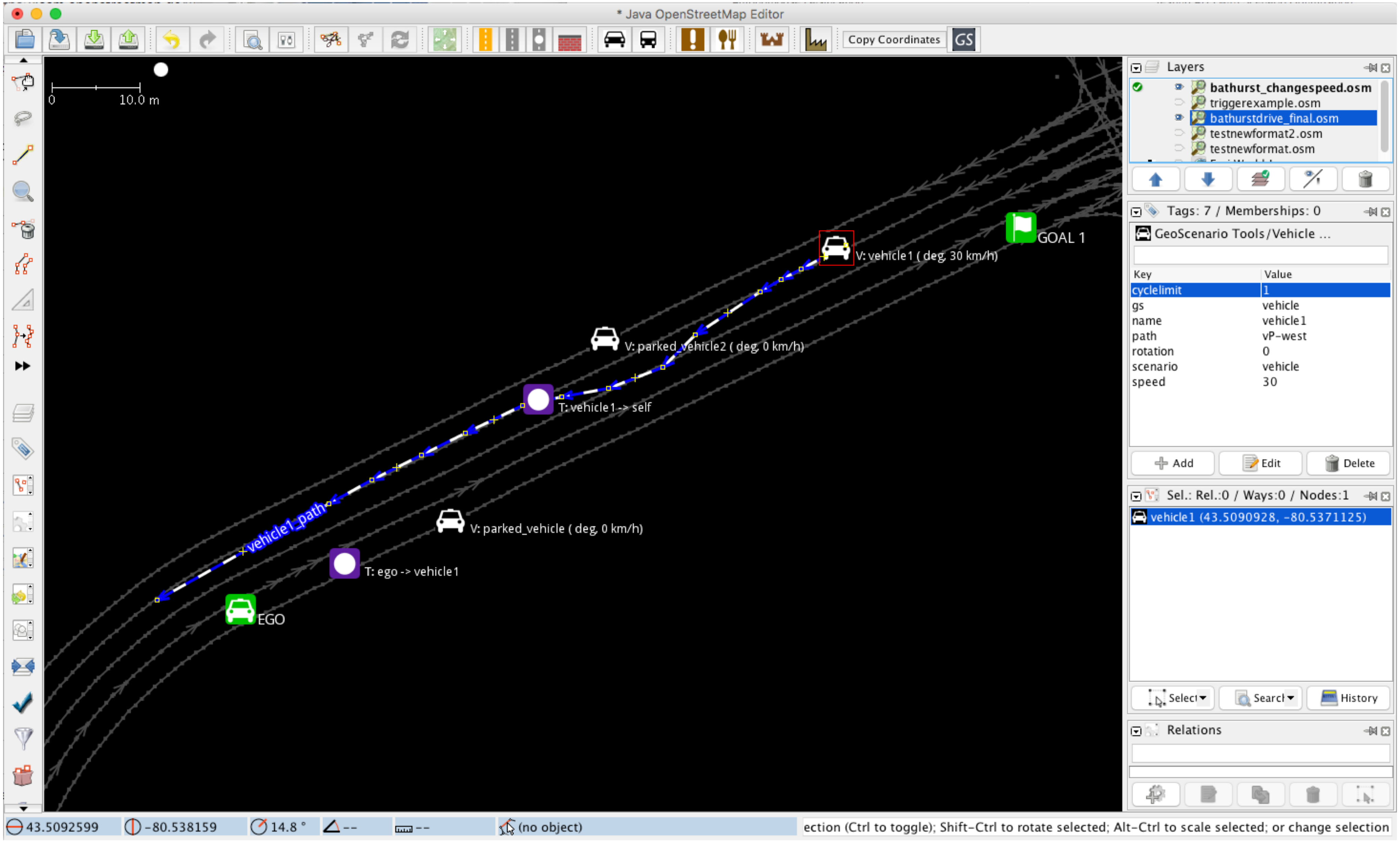
Birds-eye view

Stanford University Experiment

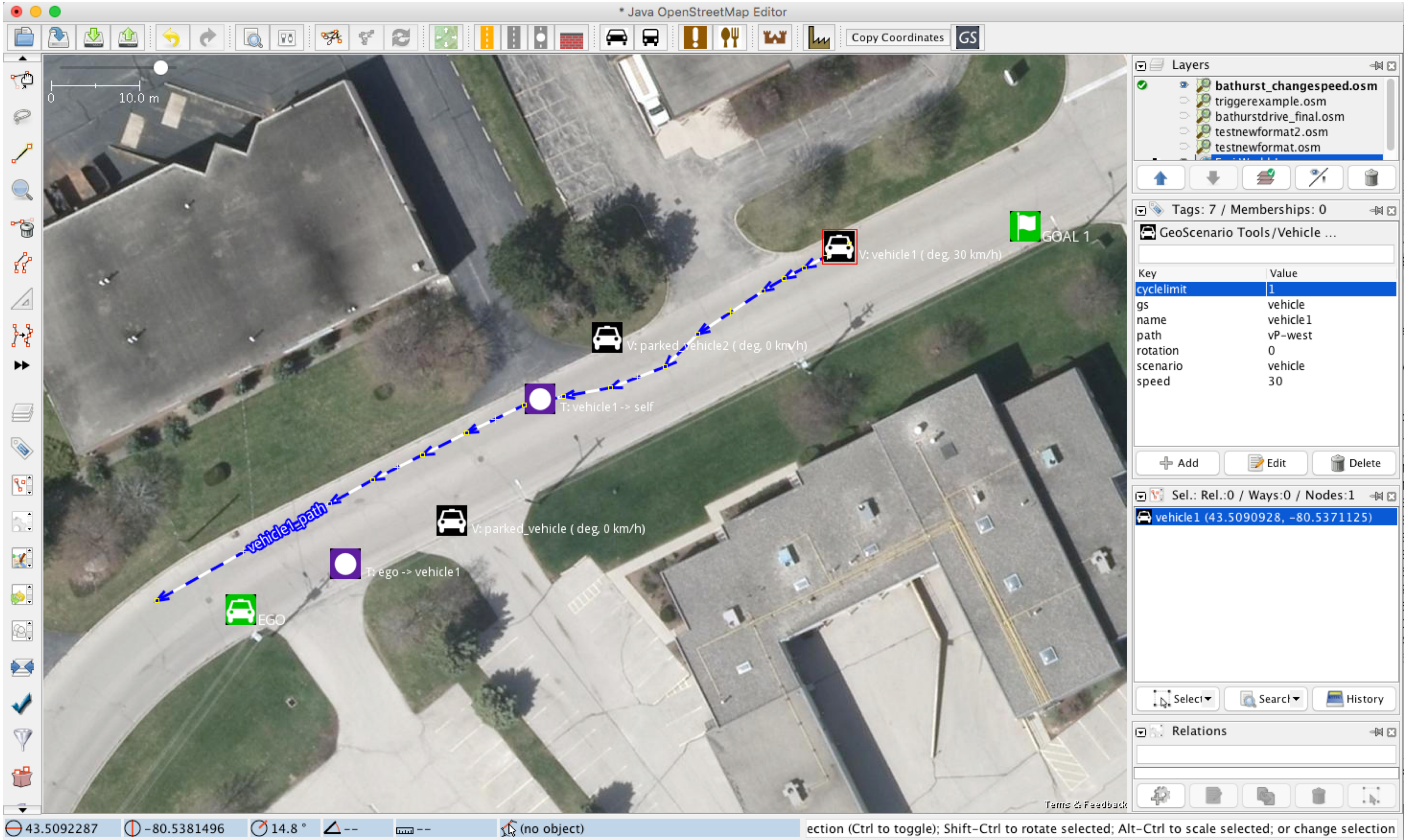
WISE Lab Simulation Environment for AV Testing

- Scenario definition in GeoScenario
  - Similar to Open Scenario
    - Location-, time-, and attribute-based triggers
  - Defined as a layer in Open Street Map
- Execution in UE4
  - Bounding box simulation
  - LIDAR simulation
  - Support for HD map
  - Collection of scoring metrics
  - Integration with ROS
  - Precise physics-based vehicle model

GeoScenario Test Definition



GeoScenario Test Definition

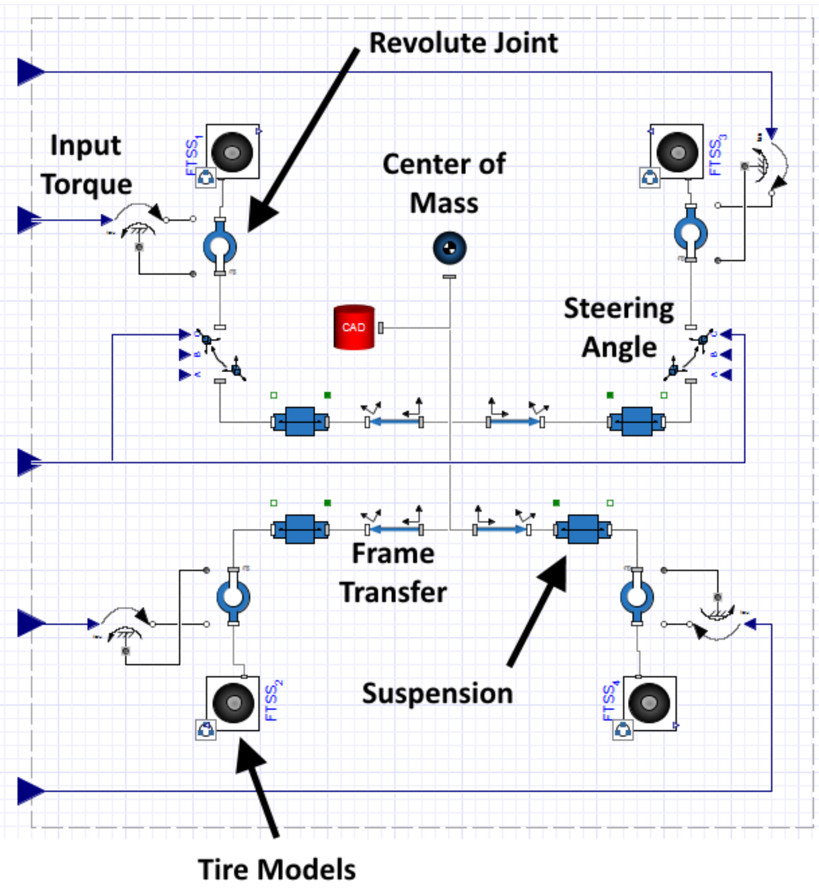


Test Execution in UE4

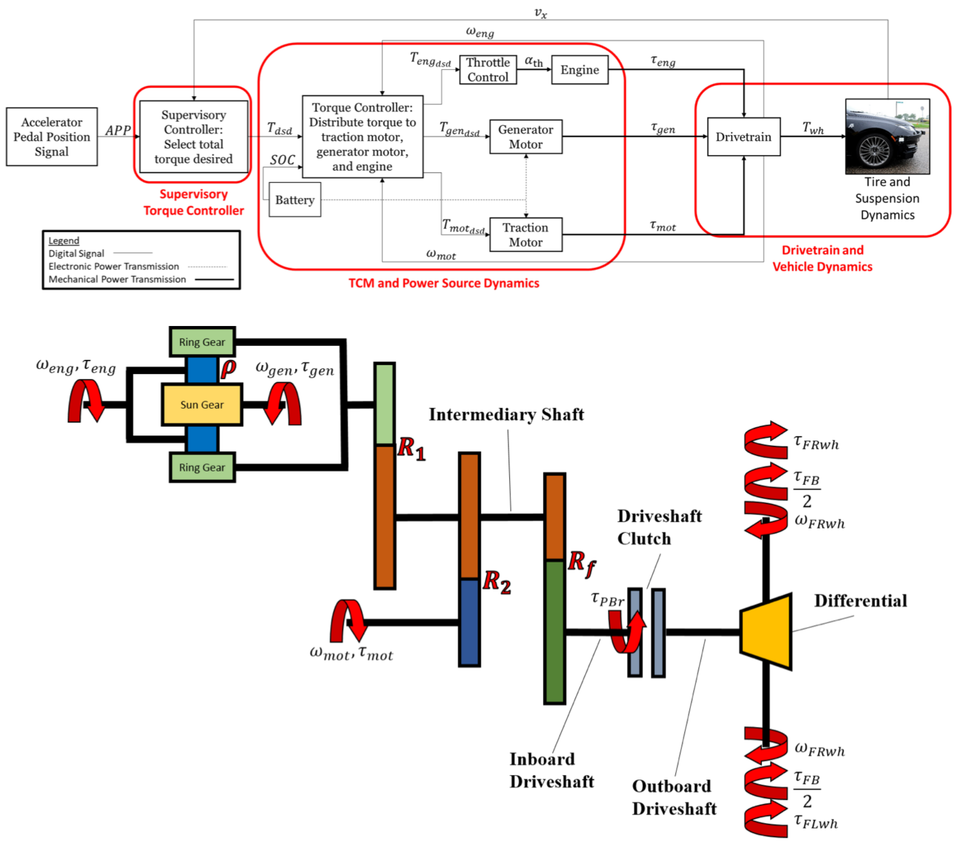
Test Execution in UE4

Vehicle System Identification

Vehicle Model in Maplesim



14 DOF vehicle dynamics model with Pacejka tires



Hybrid powertrain model
(incl. power management software)

https://uwspace.uwaterloo.ca/handle/10012/14094

Human Road User Models

https://arxiv.org/abs/1903.01539

A behavior driven approach for sampling rare event situations for autonomous vehicles.

Atrisha Sarkar and Krzysztof Czarnecki
University of Waterloo
atrisha.sarkar@uwaterloo.ca, kczarnec@gsd.uwaterloo.ca

***Abstract*— Performance evaluation of urban autonomous vehicles requires a realistic model of the behavior of other road users in the environment. Learning such models from data involves collecting naturalistic data of real-world human behavior. In many cases, acquisition of this data can be prohibitively expensive or intrusive. Additionally, the available data often contain only typical behaviors and exclude behaviors that are classified as rare events. To evaluate the performance of AV in such situations, we develop a model of traffic behavior based on the theory of bounded rationality. Based on the experiments performed on a large naturalistic driving data, we show that the developed model can be applied to estimate probability of rare events, as well as to generate new traffic situations.**

I. INTRODUCTION

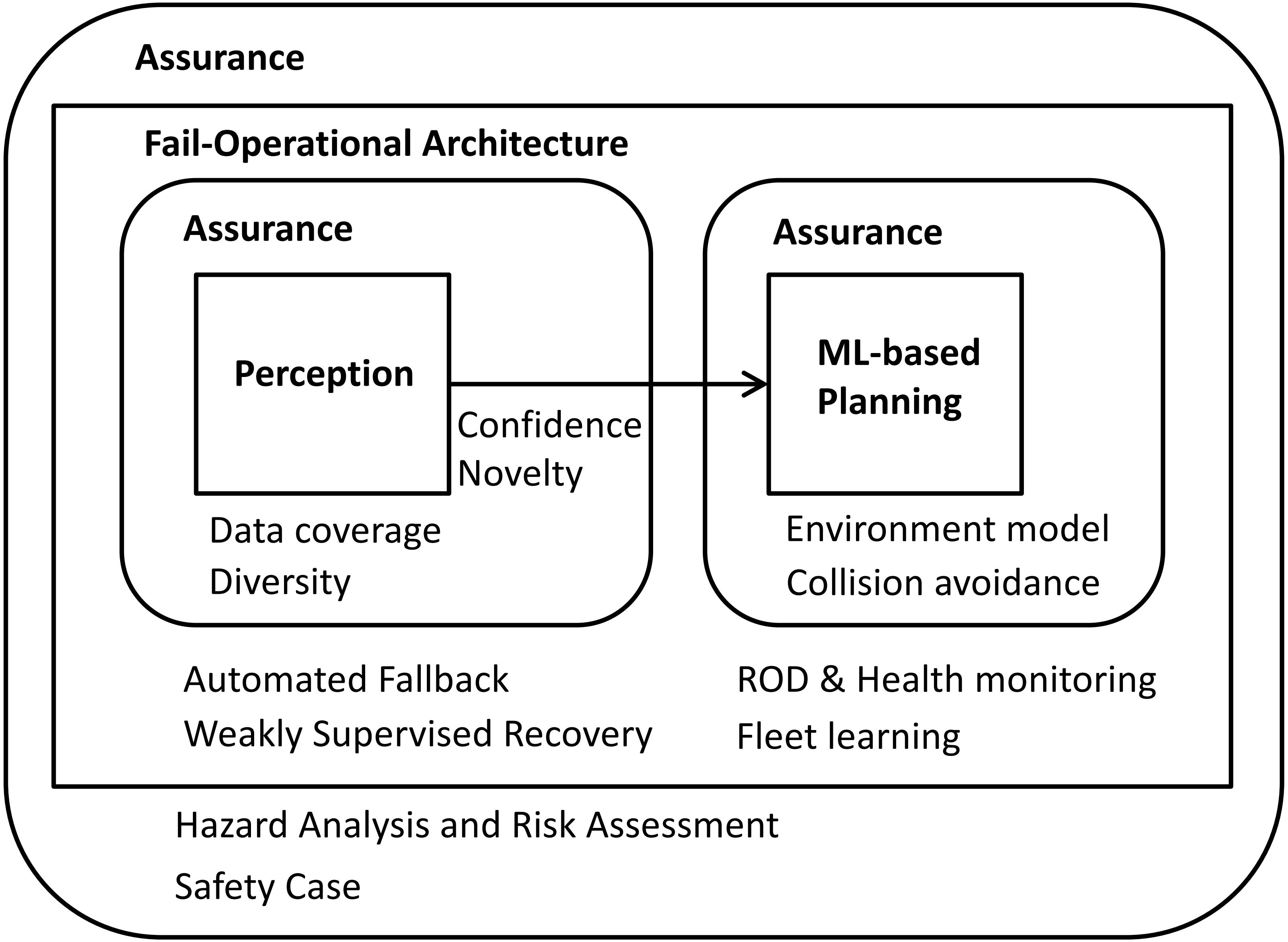
With autonomous vehicles (AV) poised to change the transportation landscape, the ability of AVs to handle a wide range of human traffic behaviors safely and reliably is of paramount importance. In order to guarantee that, it is

In recent years, RE sampling based techniques have been used for simulation based verification and testing of a wide range of motion and behavior planners. O'Kelly et al. use RE sampling for testing of planners that work in end-to-end manner based on deep learning [7], whereas, other approaches apply similar techniques to evaluate performance in specific traffic situations, such as lane changes and cut-ins [8]. Most approaches that use rare event sampling for AV evaluation, uses cross-entropy based importance sampling, which is an adaptive sampling technique to search for a sampling distribution that maximizes odds of leading to crashes and near-miss scenarios.

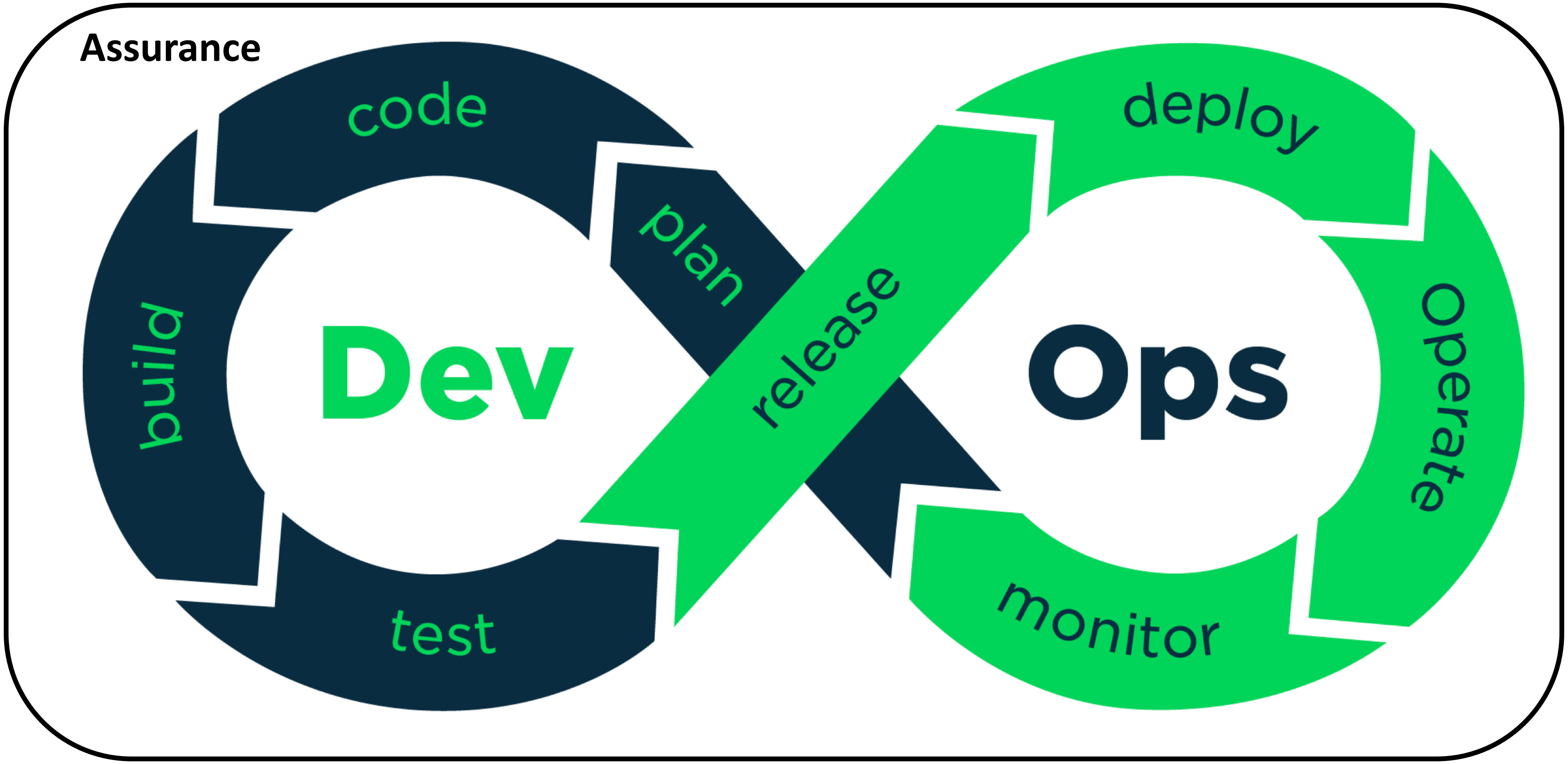
A part of the uncertainty in traffic environments arises from the inherent stochastic behavior of road users, as reflected in different driving styles of human drivers. This is in contrast to the design of motion and behavior planners

Summary

LAVA: Learned & Assured Vehicle Autonomy



DevOps for ADS Software



Shadow testing

Design of experiments & fleet learning

What field data to collect?

Update assurance

Incremental assurance

Safety case evolution