CITY
UNIVERSITY OF LONDON
EST 1894

# 89th IFIP 10.4 Meeting – Kaunas, Lithuania

## Session 3 summary

Presented by Ilir Gashi

# Overview

- Title of session: **Evolving Safety and Security in Intelligent Systems**
- Three talks:
  - **Al Avižienis**, "What motivated the concept of fault tolerance"
  - **Pascal Traverse**, "(Careful) Disruption for Safety“
  - **Wilfried Steiner**, "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

# **Al Avižienis** "What motivated the concept of fault tolerance"

- A wonderful set of slides and written note from Al describing both his personal and technical/research journey.
- The early work with JPL on the TOPS for the PGT – proposed building an **autonomous** on-board computer for TOPS named STAR (Self Testing And Repairing).
  - The two spacecrafts (named “Voyager”) were launched in 1977 and are still communicating outside of our Solar system – some of the survival features of the STAR system were included.
- Many other features of STAR were adapted and developed further in subsequent decades – supported initially by a five-year grant from NSF in transferring the research from NASA to UCLA, and subsequent support from NSF, NASA, DoD, FAA and industry – the first new major project was the experimental investigation of “N-version programming”.
  - 10 faculty members, 20 visiting scholars, and 50 gradaute students took part on the research directed by Al Avizienis.
- Al concluded his talk with a proposal to the working group to change the name of the WG to “Working Group 10.4 on Dependable and Trustworthy Computing”.

# **Pascal Traverse**, "(Careful) Disruption for Safety“

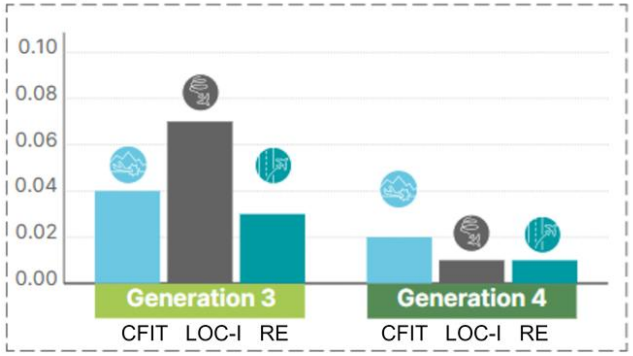
- Fly-by-Wire technology is adopted by all aircraft manufacturers and FbW airplanes have a good safety record.
  - Its introduction, which relies on a new concept of operations and the use of software, **was seen as disruptive**, even if it was carefully done.
- The talk explored
  - A new step in cockpit automation
  - The use of Artificial Intelligence.
- Both cases were discussed from three view points:
  - the potential for safety increase,
  - the precaution to be taken, and
  - if it would be disruptive.

# "(Careful) Disruption for Safety“

## (Careful) Disruption for Safety

## Safety?

Average fatal accident rate

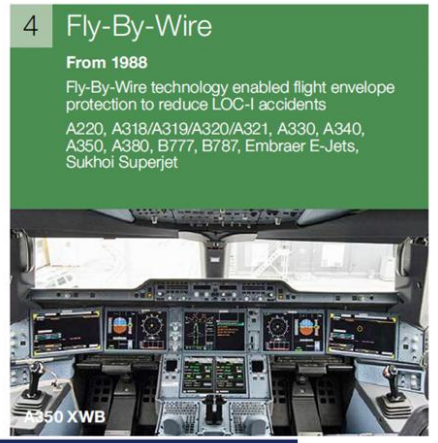



Average fatal accident rate (per million flights) per accident category 1958-2025

- Controlled Flight Into Terrain
- Loss Of Control In flight
- Runway Excursion

<= Flight Envelope Protection (LOC-I)

## A320 Fly-by-Wire


# "(Careful) Disruption for Safety“

(Careful) Disruption for Safety

Disruptive Cockpit - DISCO

## Safety?


1987 ------ 1993 ------ 2002

- Flight envelope protection
- Terrain Awareness & Warning System
- Traffic Collision Avoidance System


2005

- Auto Pilot - TCAS
- Runway Overrun Protection System


2013 ------ 2020 ------ 2024

- Auto Emergency Descent
- Robust Autopilot

## Next?

*=> Give time and data to the crew, in all conditions, to assess the situation and make a good decision.*

# "(Careful) Disruption for Safety“

(Careful) Disruption for Safety

Disruptive Cockpit - DISCO

What is it about?

*=> Give time and data to the crew...*

***Automation***
***Perception***
***Navigation***
***Human-Machine Interface***

# "(Careful) Disruption for Safety“

(Careful) Disruption for Safety

Disruptive Cockpit - DISCO

## Careful?

- Independent assessment panel
- Feedback from airline pilots
  DISCODECK
- Integrate the technologies
  DISCOBENCH
- Fly the critical technologies
  ATTOL - take-off and landing
  DRAGONFLY - continued safe flight and landing
  OPTIMATE - taxi

# "(Careful) Disruption for Safety“

(Careful) Disruption for Safety

Flying AI

| ***Advisor*** | ***Virtual Co-Pilot*** | ***Savior*** |
|---|---|---|
| Frequent Support | Rare Support | Exceptional Support |
| => ATC messages<br>=> document search<br>=> trajectory | => detection of incapacitation<br>=> continued safe flight and landing | => invention of a solution |

# "(Careful) Disruption for Safety“

(Careful) Disruption for Safety

Flying AI

European Union Aviation Safety Agency

**Notice of Proposed Amendment 2025-07 (B)**

in accordance with Article 6 of Management Board Decision 01-2022

| ***Advisor*** | ***Virtual Co-Pilot*** | ***Savior*** |
|---|---|---|
| Frequent Support | Rare Support | Exceptional Support |
| => No stress<br>=> “Major” | => Remote Activation<br>=> Limitations<br>=> Vote? | => Extremely Remote Stressed Crew Activation |
| DAL C? | DAL B? | DAL B? |

# **Wilfried Steiner**, "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

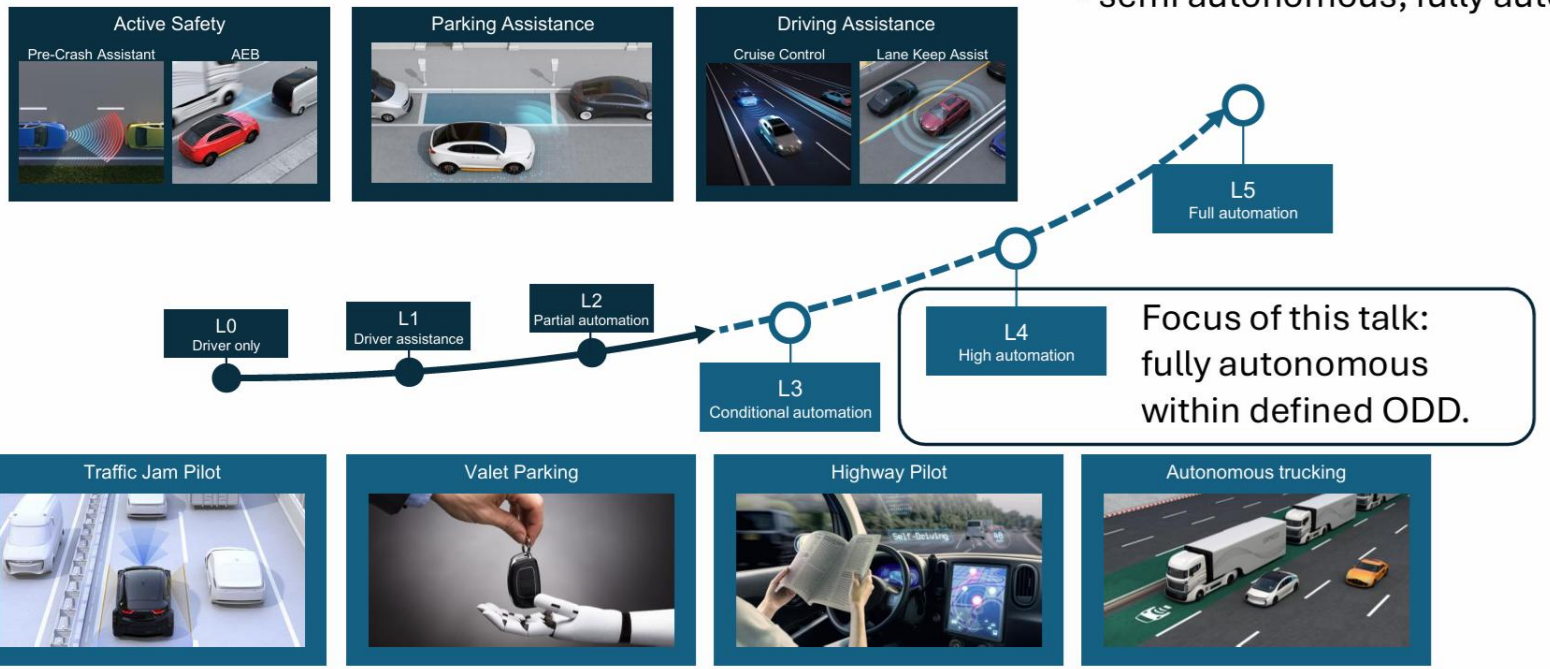
Today: Cars have not sufficiently been shown to be truly self-driving even within limited operational design domains (ODDs), revealing a remaining autonomy gap.

Tomorrow: While this autonomy gap is being closed, research should expand toward large-scale Systems of Autonomous Systems (SoAS), including self-driving cars and extending beyond them to other domains.

# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## SAE Levels of Autonomy

Alternative classifications have been proposed, e.g.,
- hands-on/off, eyes-on/off (3 cases)
- semi autonomous, fully autonomous

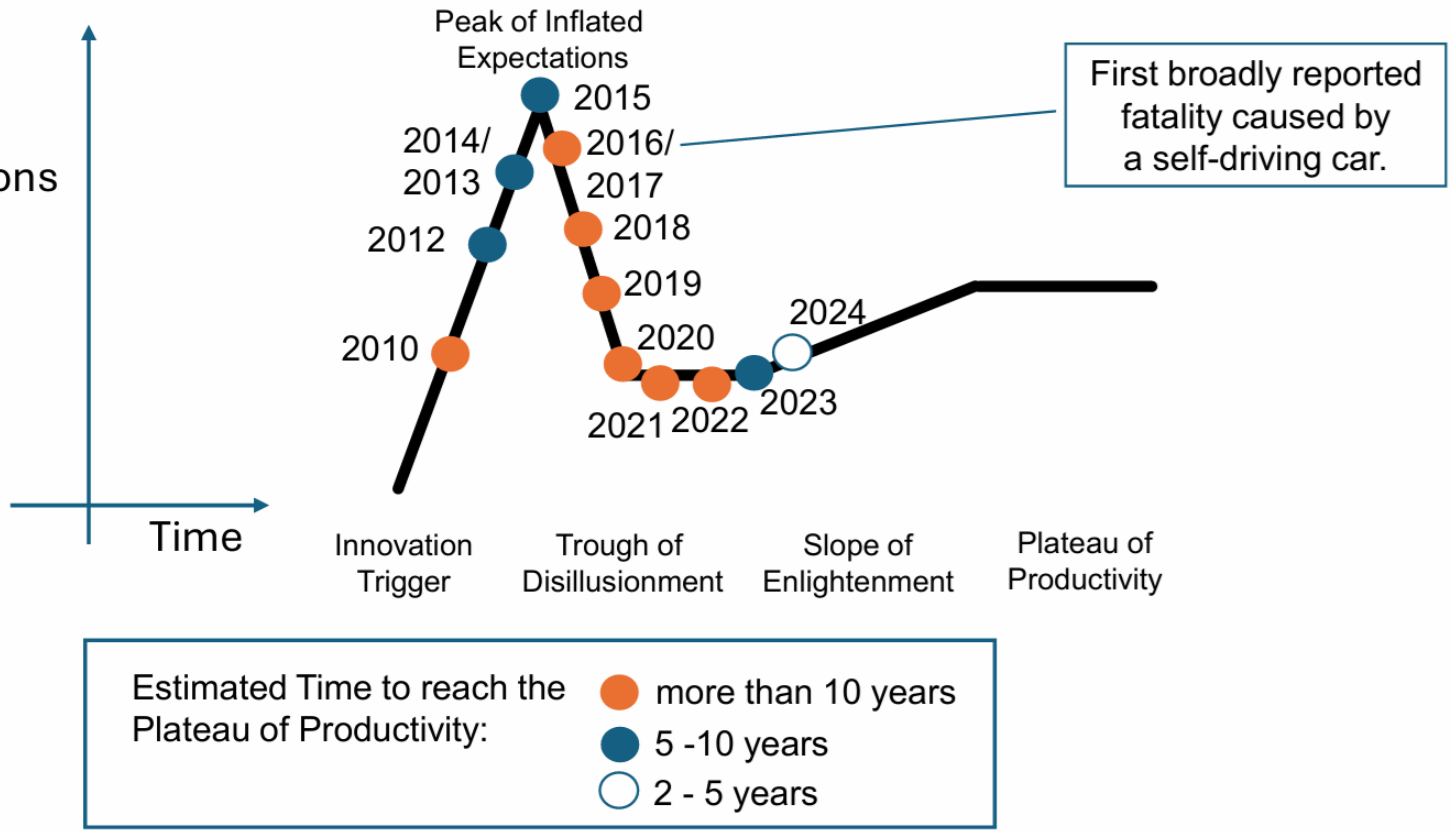



ODD ... Operational Design Domain

# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## Current state of Self-Driving Cars

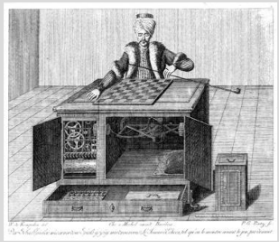
Data from Gartner's Hype Cycles 2010 - 2024



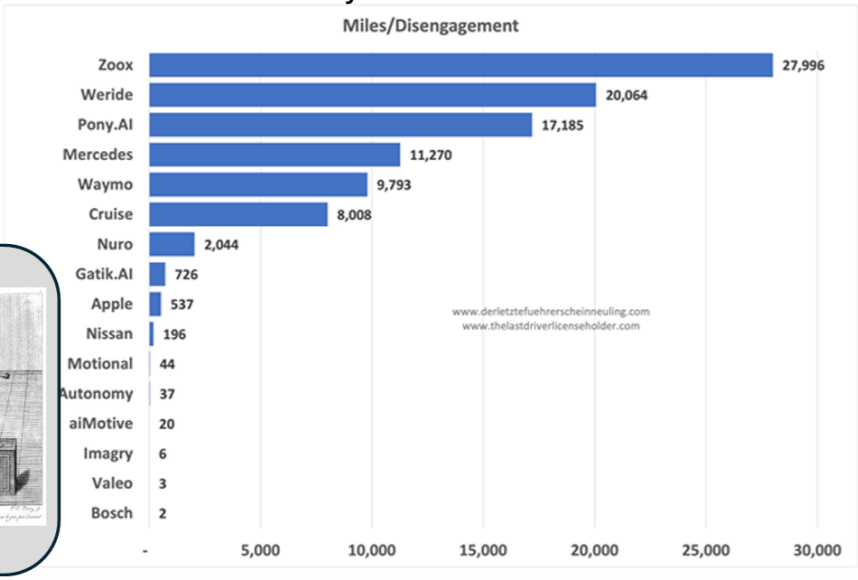
# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## How "self-driving" are cars today?

- Tele-operator
  - A human outside the vehicle remote controls the vehicle.
- Safety driver
  - A human fallback inside the vehicle intervenes in critical situations.
- Trail car
  - A human safety driver operates from a separate vehicle.
- Tele-assistance
  - A human assists the vehicle, upon vehicle's request.



Average miles driven before human safety driver intervention:

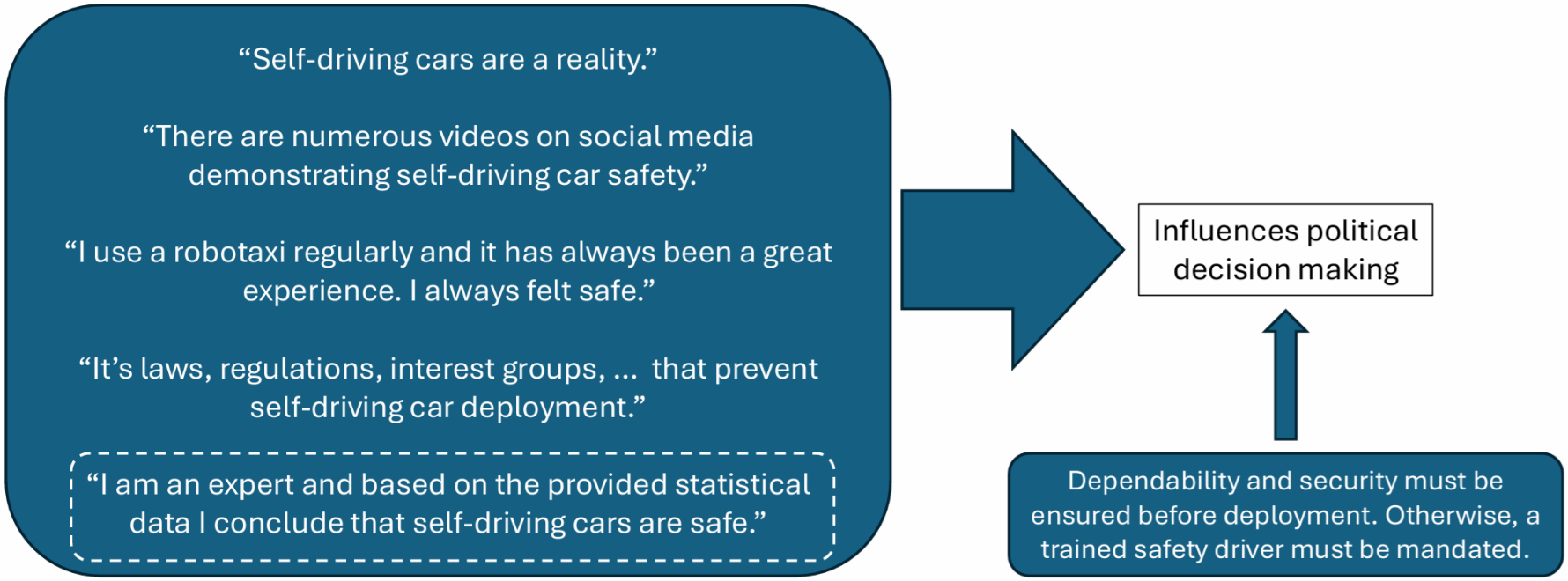



2024 Disengagement Reports from California

*) The computer-based driving system "disengages".

# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## Self-driving cars and the public opinion (simplification)



# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

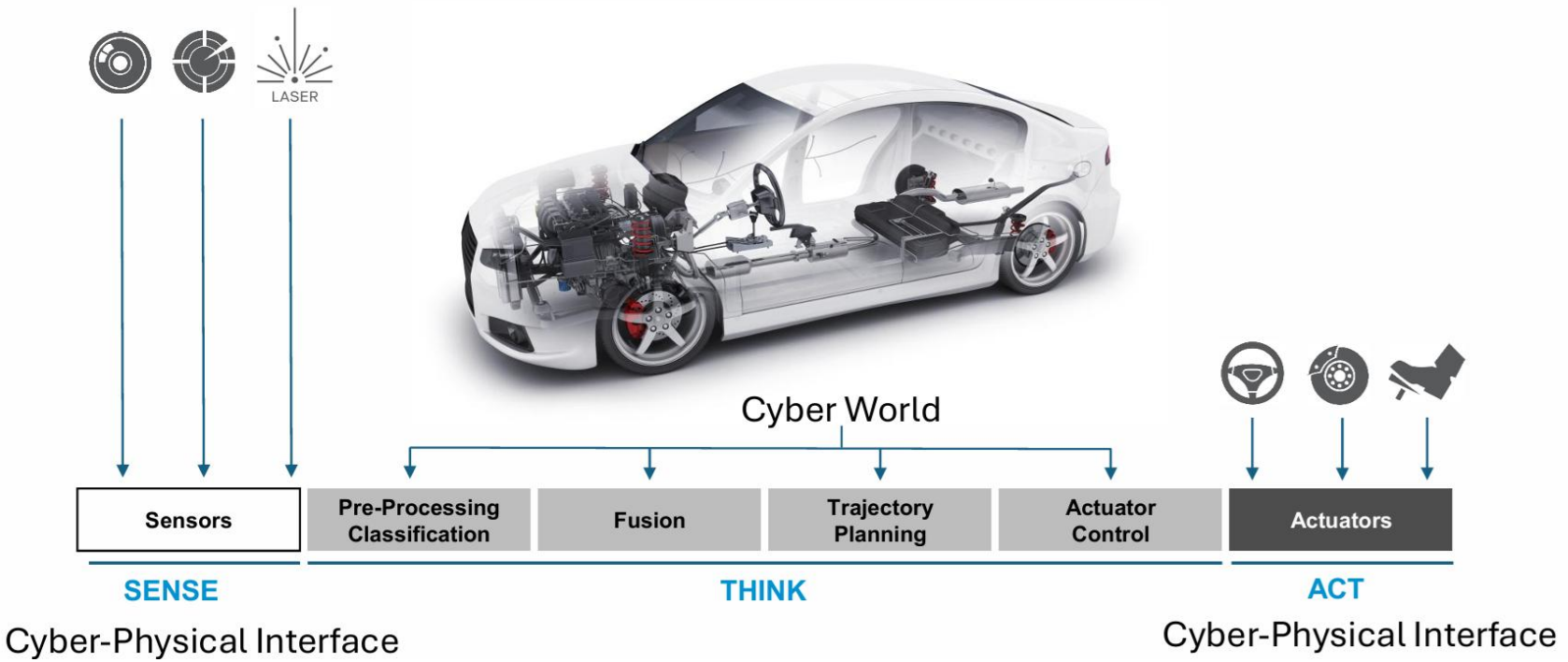
## Near-term challenges (2026-30)

- The Intelligent Vehicle itself
- The infrastructure that the Intelligent Vehicle is relying on
- New technologies in development and design
- Legislation and policy

# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

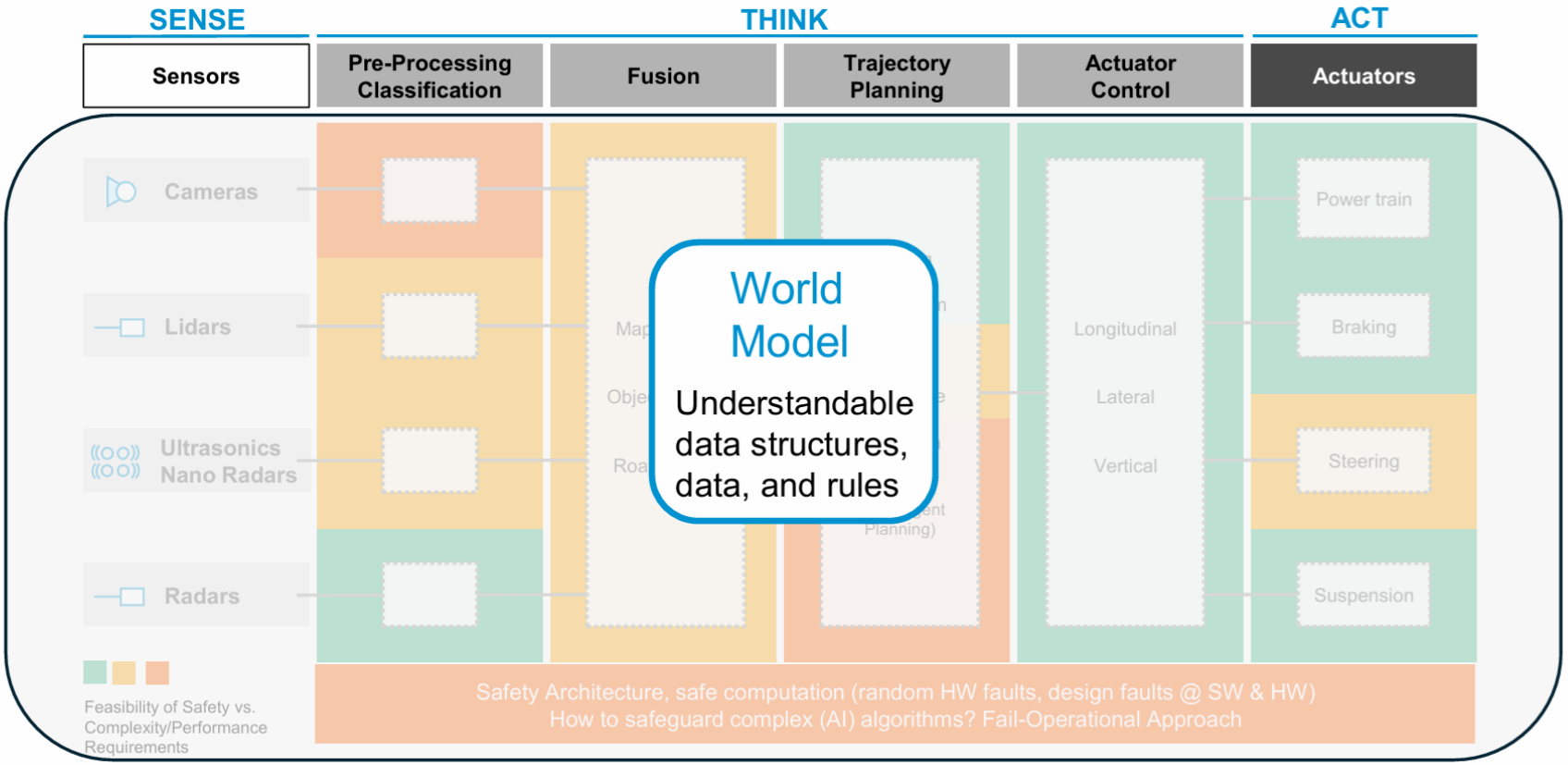
## Challenges regarding the Intelligent Vehicle itself

- Cyber World, Physical-World, and the Cyber-Physical Interface



# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

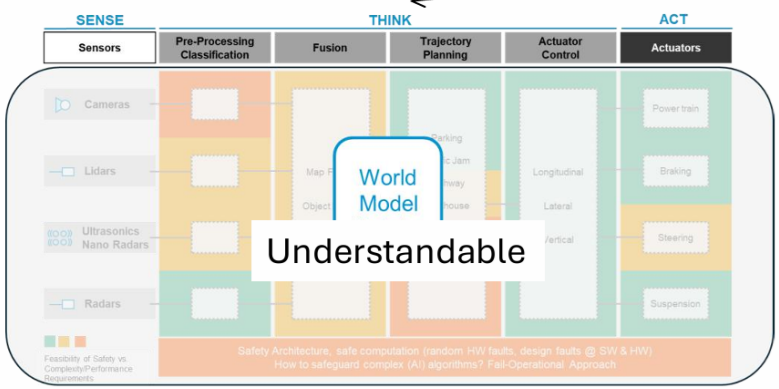
## The Importance of the World Model



# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## Two Approaches:
## Distributed Decomposable Architecture vs. Monolithic End-to-End

This must run on a fault-tolerant computer.

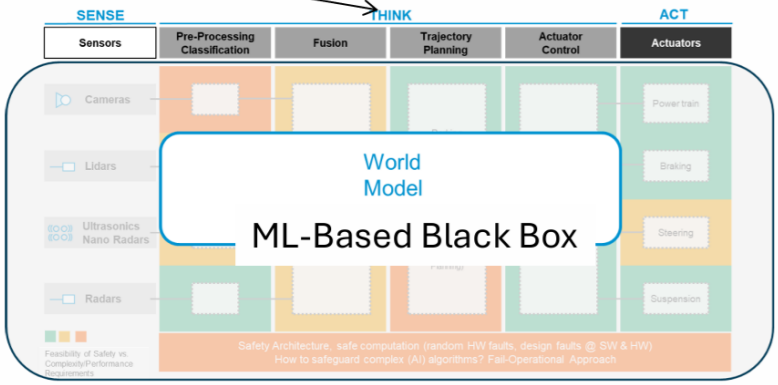



Distributed Decomposable Architecture

Likely the better choice for safety and security assurance.

Enables Continuous Validation, e.g., using predictive processing*

- at t1: compute model state at t2
- at t2: check if sensor readings confirm computed state




Monolithic End-to-End

Restricts V&V possibilities to the full system only.

* Rushby, J. (2022). *Models and their Validation and their Role in Perception And in Safe Autonomous Vehicles*. IFIP WG10.4 Virtual Meeting.
URL: http://www.csl.sri.com/users/rushby/abstracts/ifip-11may22

ifip WG.10.4

# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## Challenges regarding the Intelligent Vehicles itself

Top-Down Design of a Decomposable Architecture

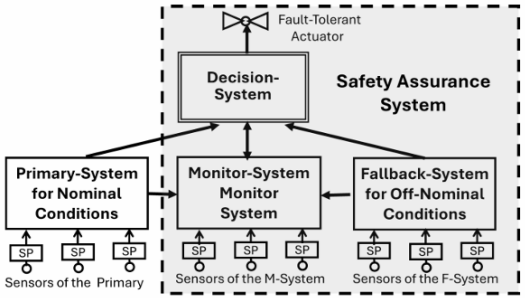



Figure depicts "logical" units. Ideally, they would be implemented as separate & diverse ECUs.

Mapping to a system of ECUs, under constraints

Design Constraints (examples):

- Legacy re-use (full ECUs, technologies, etc.),
- SWaP-C,
- HW performance,
- Availability of certifiable HW & SW components,
- Scalable to support model variants

E/E Architecture options

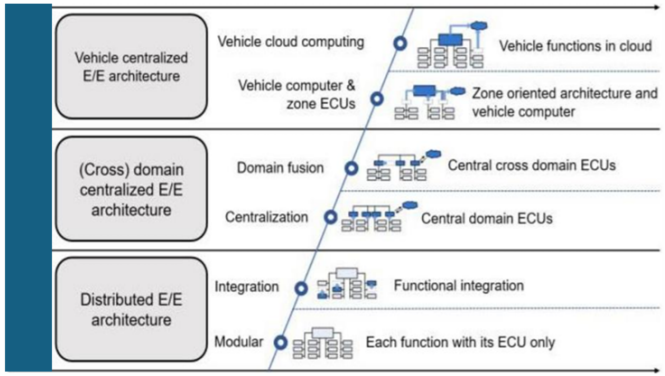
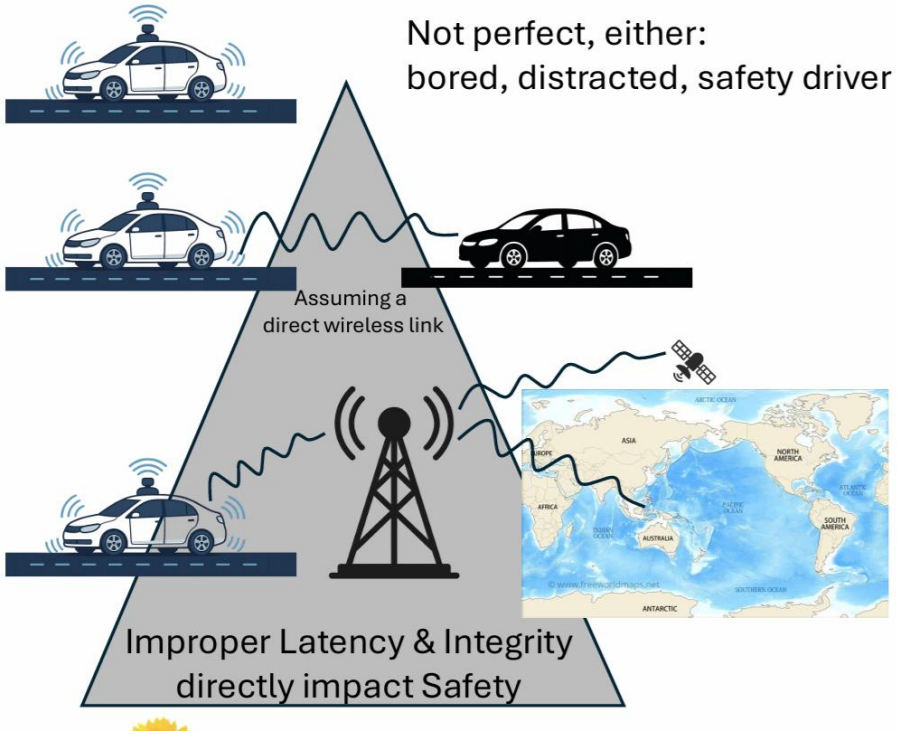



Figure depicts different ECU arrangement strategies.

# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## Challenges regarding the infrastructure that the Intelligent Vehicle is relying on

- Safety driver
  - A human fallback inside the vehicle intervenes in critical situations.
- Trail car
  - A human safety driver operates from a separate vehicle.
- Tele-assistance
  - A human assists the vehicle upon vehicle's request.
- Tele-operator
  - A human outside the vehicle remote controls the vehicle.



# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## Challenges regarding new technologies in development and design

Positive example of GenAI use:

- KitKat experiment*): LLMs list and explain dependability issues of the KitKat incident

Negative example of GenAI use:

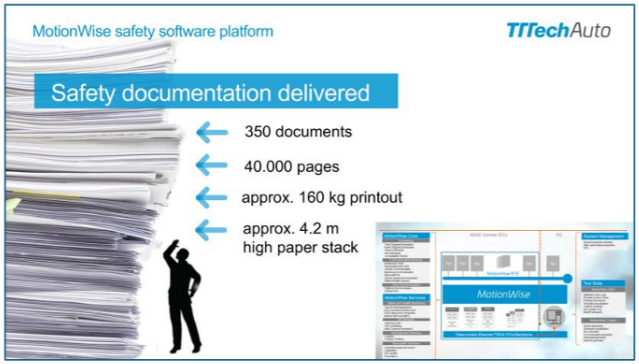
- Stadler Metro Car experiment*): LLMs wrongly assign new cars to different target customers, even when iteratively feeding them the respective other answers

*) see annex slides for experiment details

Continuous improvements and role of neuro-symbolic AI (symbiosis of ML and formal methods).

Still: for now, use Zero Trust methodology when dealing with LLMs. Never trust, always verify.

Opportunity: using LLM technology for development and design

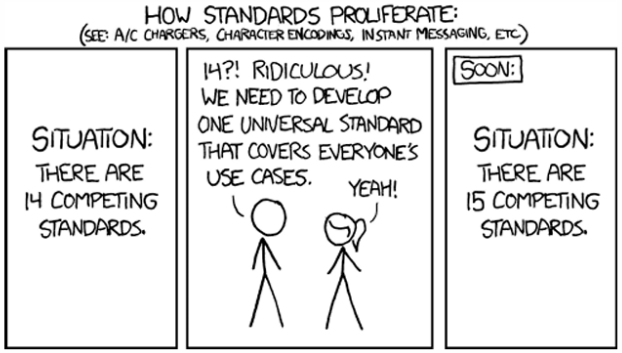



Typical traditional example. How much of such artefacts could be generated by GenAI?

# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## Legislation and policy challenges

- Availability and transparency of field data regarding incidents and tele-assistant help of vehicles on the road today
  - While California collects some data, not all states where driverless-cars are allowed to experiment do so.
  - CA data is a good start but is not comprehensive enough to assess state of current operational safety and security.
  - It also does not shed light on how frequently tele-assistants have to help and what situations lead to self-driving systems being stuck.
- Fragmentation and localization of regulations and laws is problematic




https://xkcd.com/927

# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## Possible Role of IFIP WG 10.4 Today (2026-30)

- Develop quantitative acceptance criteria for solutions that use AI/ML.
  - (#miles driven is the insufficient metric – how many edge cases were encountered, and the vehicle response are more informative)
- Definition of the means of deciding whether the acceptance criteria are met.
  - Develop guidelines for collecting field data regarding incidents and tele-assistant help.
- Develop methods to construct the overall system assurance case.
- Give recommendations for preferred technologies to realize IVs and their usage that ease the assurance case argumentation.
  + Develop technologies that ease the assurance case argumentation.

# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## Longer-term challenges (2030-40)

- Regulatory bodies establish safety and security standards.
- AI/ML algorithms can classify objects with extremely high level of fidelity.
- Operational Design Domain (ODD) Extension: Testing, validation, and verification methodologies can account for extreme cases of operational environment, including weather, obstacles, people, other vehicles, and emergency activities.
- **Impact of non-availability of fleets of autonomous systems.**

# "Intelligent Vehicle Dependability and Security: Challenges for Today (2026-30)

## Possible Role of IFIP WG 10.4 Tomorrow (2030-40)

*) Assuming positive development of self-driving cars / robotaxis.

- Research dependability and security of a System-of-Autonomous-Systems that include:
  - fleets of self-driving cars,
  - autonomous duty and transportation vehicles,
  - infrastructure (including maintenance and construction machines),
  - unmanned aerial vehicles and vertical take-off and landing aircrafts,
  - etc.
- Address the implied scaling challenges in:
  - definition of quantitative acceptance criteria and their measurement
  - assurance case construction & system technology recommendations
  - etc.

# Q&A

- Good Q&A on the talks
  - The effect on certification from moving to Single-Pilot Operation
  - Distributed Decomposable Architecture vs. Monolithic End-to-End
    - A lively discuss on what is “better” / ”leads to higher confidence that system is dependable and secure”
  - Discussion on whether to change the name of the WG

# Thank you!

- Correction/editions/clarifications are welcome (from presenters and audience).