جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

Resilient Computing and Cybersecurity Center

Resilient Computing and Cybersecurity Center

Safe and Secure AI/ML-driven Autonomous Vehicles? Not anywhere near yet …

https://rc3.kaust.edu.sa

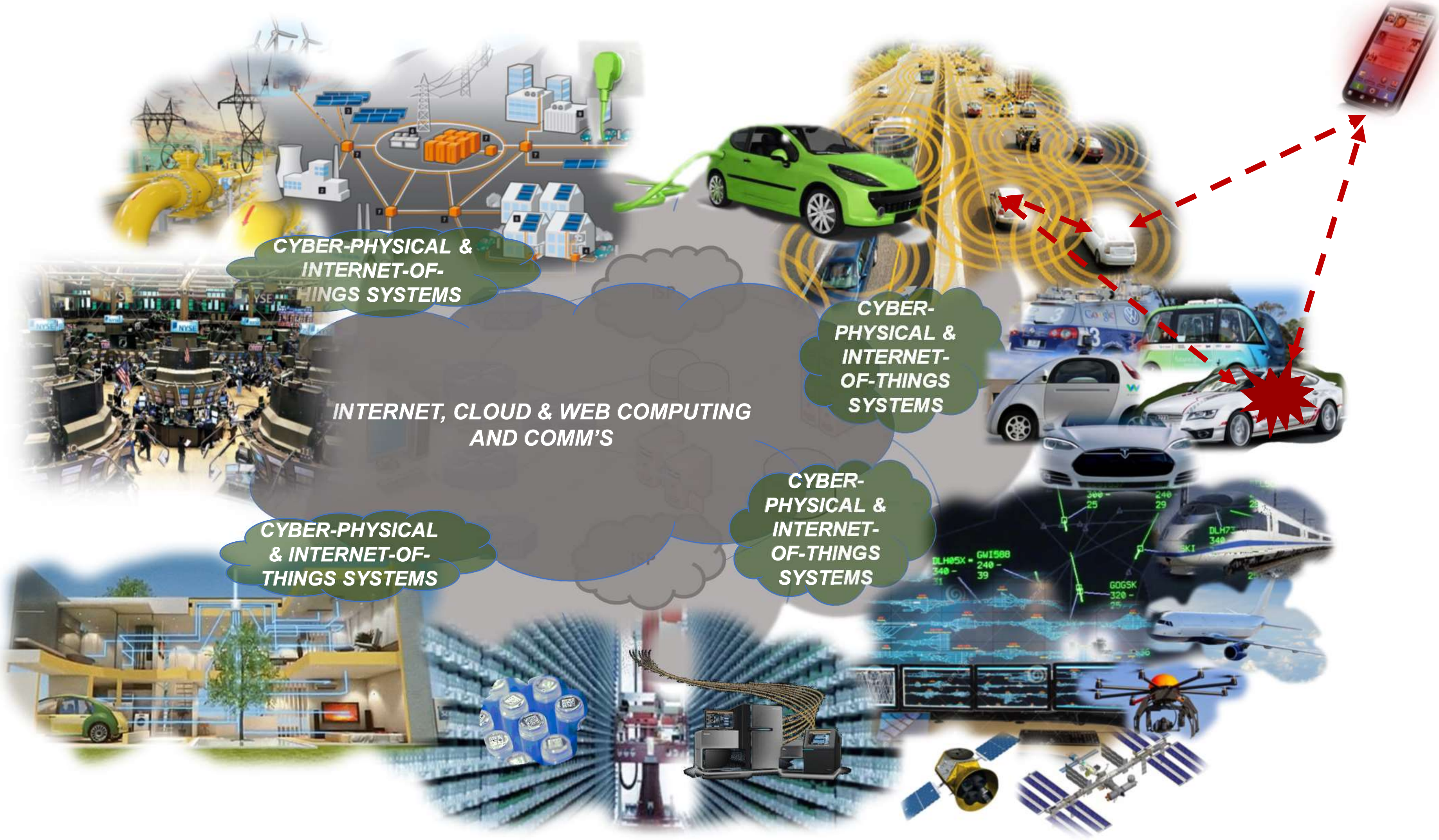
Paulo Esteves-Veríssimo, **Professor, Director**
King Abdullah University of Science and Technology, KAUST
Resilient Computing and Cybersecurity Center – RC3
paulo.verissimo@kaust.edu.sa
https://cemse.kaust.edu.sa/people/person/paulo-verissimo

rc3.kaust.edu.sa

The world *BECAME* an immense, interconnected, infrastructure

The world is becoming an immense, interconnected, infrastructure

The world is becoming
an immense, interconnected, infrastructure

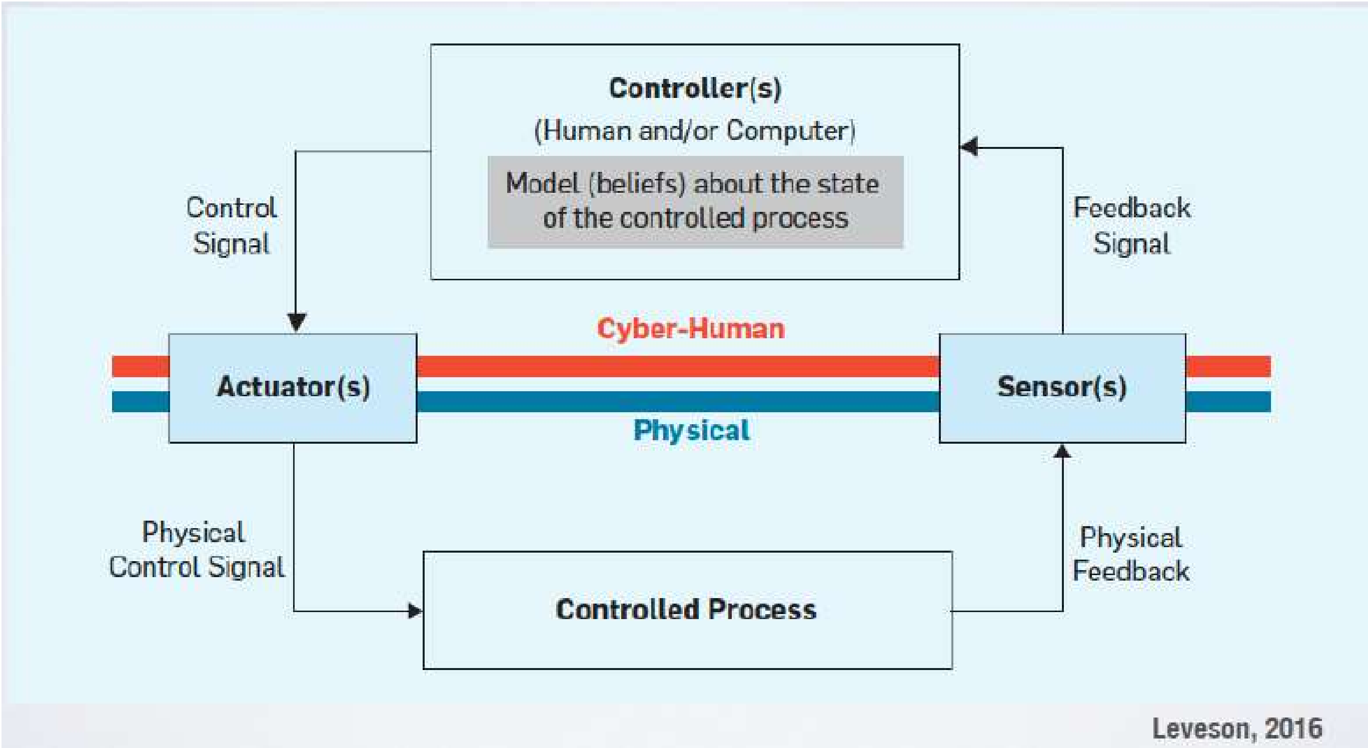


Brief Analysis of the Cyberspace *today*

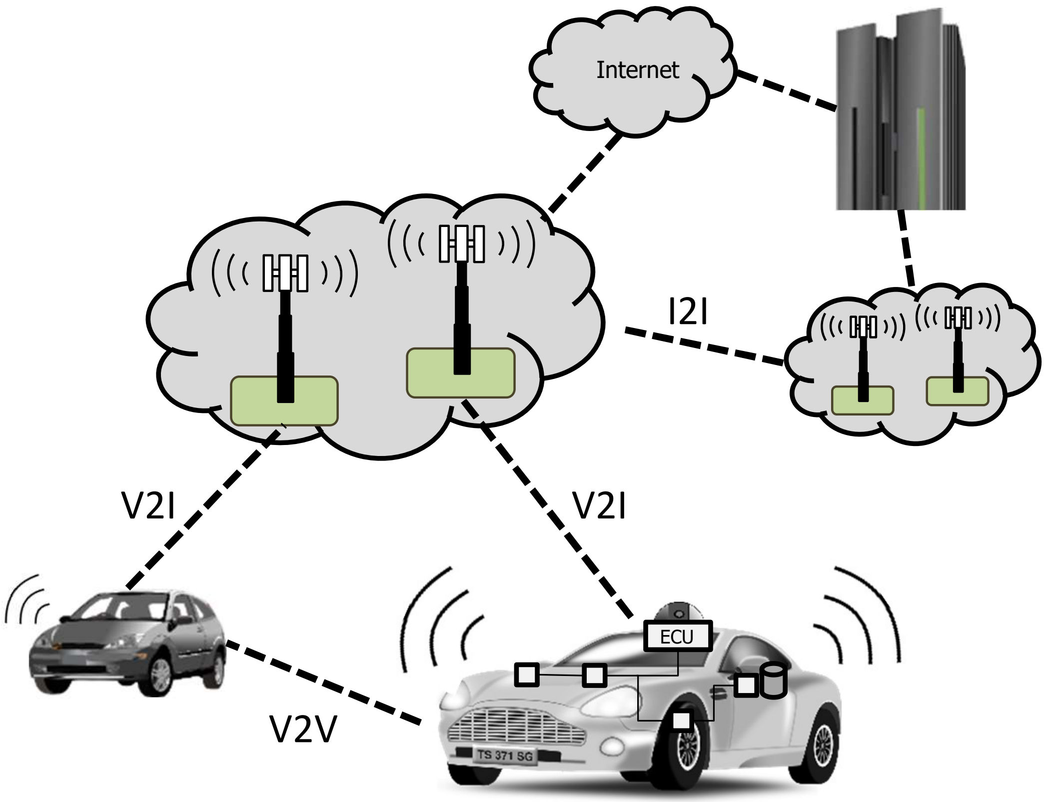
- distributed infrastructure:
  - *Pervasive CPS and IoT*; seamless integration with Internet/Cloud/Web.
- highly exposed to threats:
  - Huge *pressure to go "digital"*: Govs; BigTechs; Social nets.
- steadily increasing software vulnerabilities:
  - Common SW yearly *rate increased* 2-3-fold; *CPS/IoT* in great increase
- degradation of the threat surface:
  - *Even more* powerful adversary actors and sophisticated exploit tools

On
The cool world of autonomous vehicles

The problem of vehicle control



Autonomous Vehicle Ecosystem

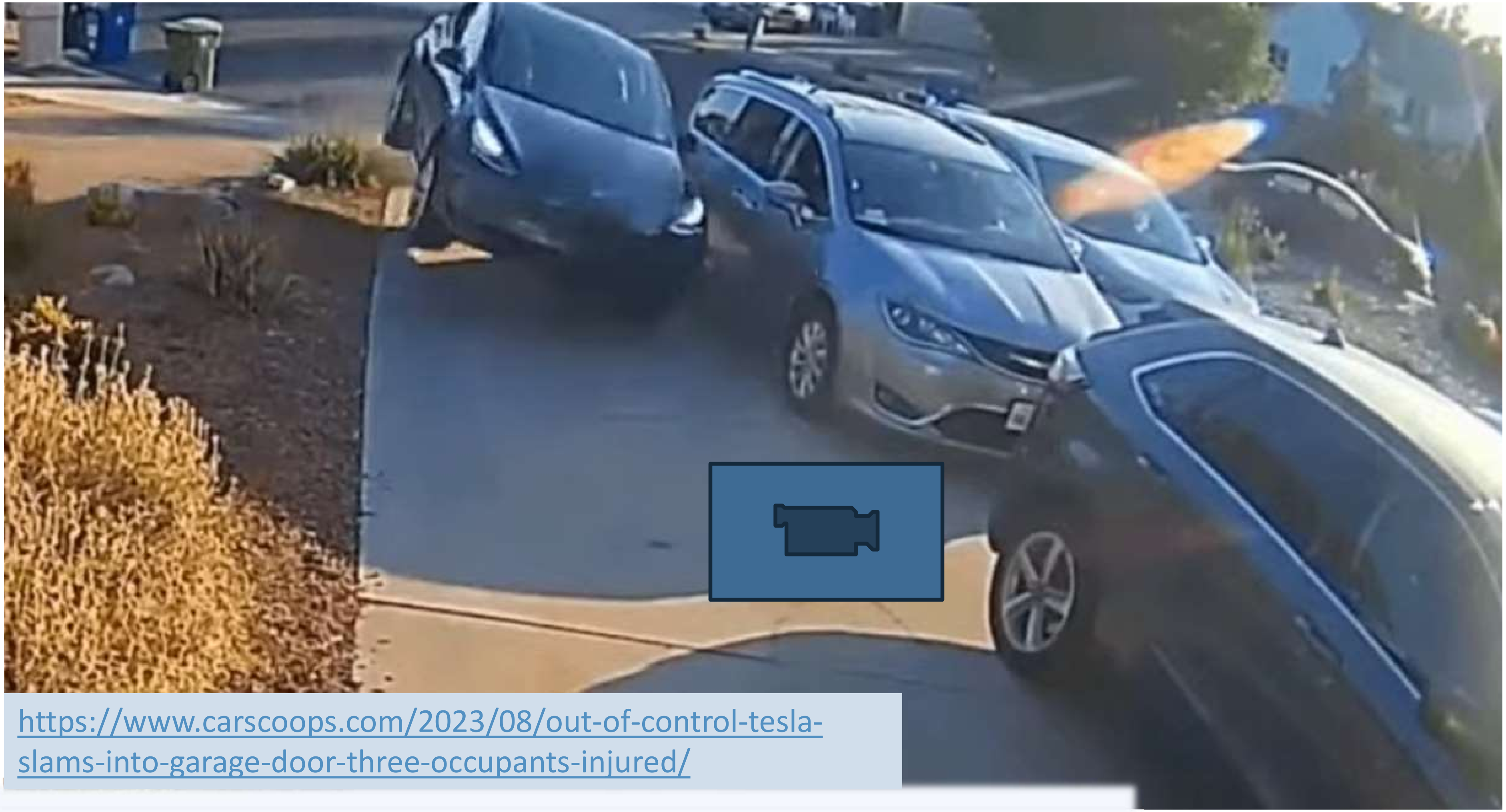


Towards Safe and Secure Autonomous and Cooperative Vehicle Ecosystems. Lima, A; Rocha, F; Volp, M; Verissimo, P. in Proc's 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy (2016, October) @CCS, Vienna-Austria

Is the *autonomous vehicles world* (cyber)-safe?

Clouds in the horizon of ... the safety side ...

It can get really bad…
BAD as in ‘out of control’



https://www.carscoops.com/2023/08/out-of-control-tesla-slams-into-garage-door-three-occupants-injured/

It can get really bad...
BAD as in 'blind'

https://www.reddit.com/r/ThatsInsane/comments/r3fxpi/tesla_radar_did_not_recognize_a_camel_cusing_an/?rdt=49822

Tesla vision did not recognize a camel, causing an accident in the UAE

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

Is the *autonomous vehicles world at least* (cyber)-dependable?

SO, what about "normal" case behaviour ? ...

King Abdullah University of Science and Technology

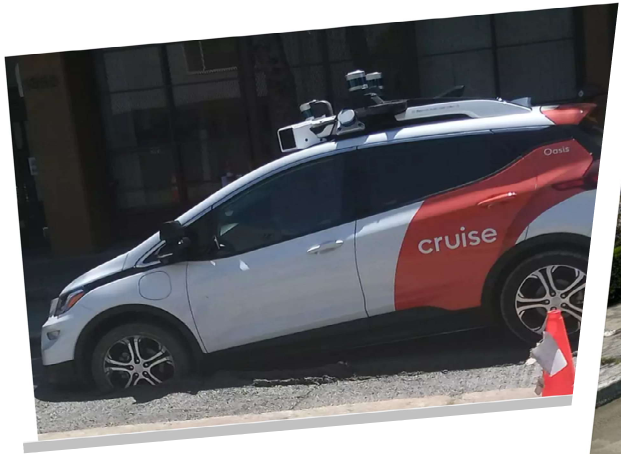
جامعة الملك عبدالله للعلوم والتقنية

SFGATE

Newsletters

Cruise vehicle gets stuck in wet concrete while driving in San Francisco

By Joshua Bote
Aug 15, 2023



r/shittyrobots • 2 yr. ago
by Jasaj4

Join ...

Delivery robot tries to walk across undried cement

CARSCOOPS

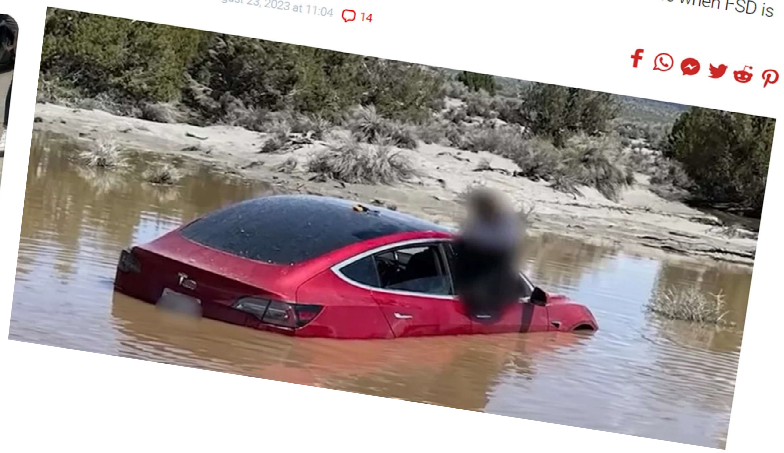
LATEST NEW CARS SCOOPS

VIDEO

Tesla Model 3 Driver Ignores FSD Limitations, Drives Through Flooded Road

The driver of the Tesla Model 3 seems to have forgotten that drivers are still responsible when FSD is

by Brad Anderson August 23, 2023 at 11:04 14



Is the *autonomous vehicles world* (cyber)-secure?

Security gap in Vehicle Systems

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

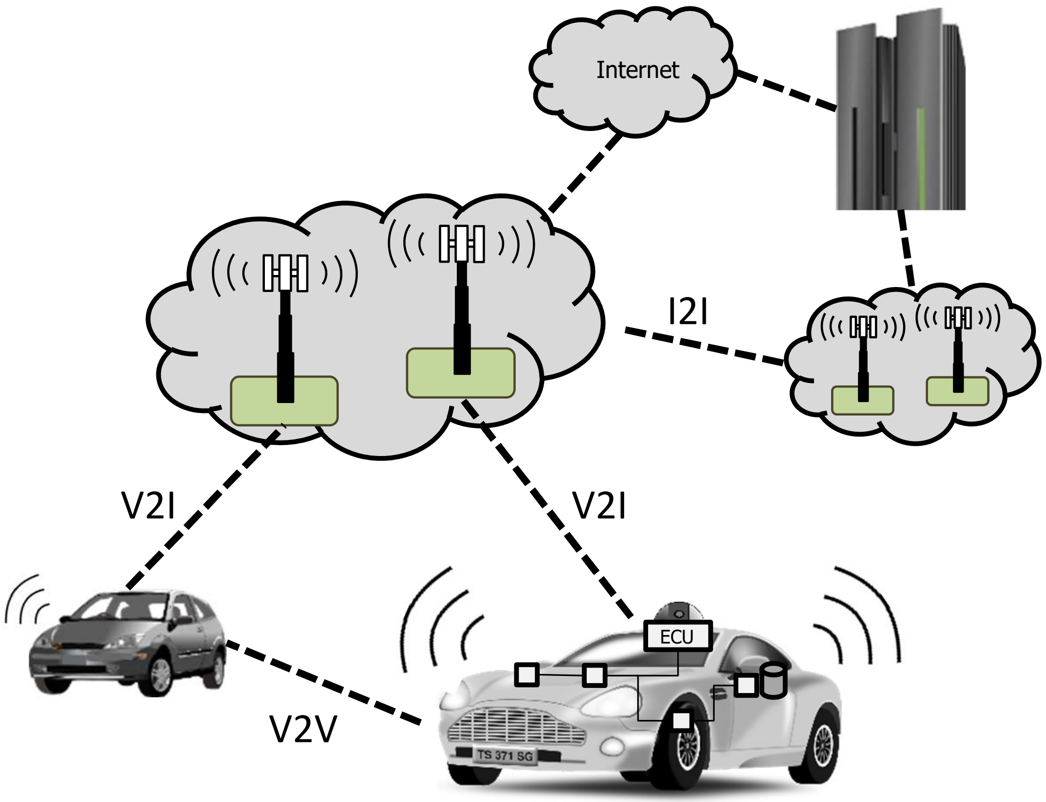
So, what's wrong about the current autonomous vehicles ecosystem?

- ***To start with, the very notion that there is an ecosystem is inexistent***

- ***An analysis of the ecosystem as a critical infrastructure is missing***

Autonomous Vehicle Ecosystem

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

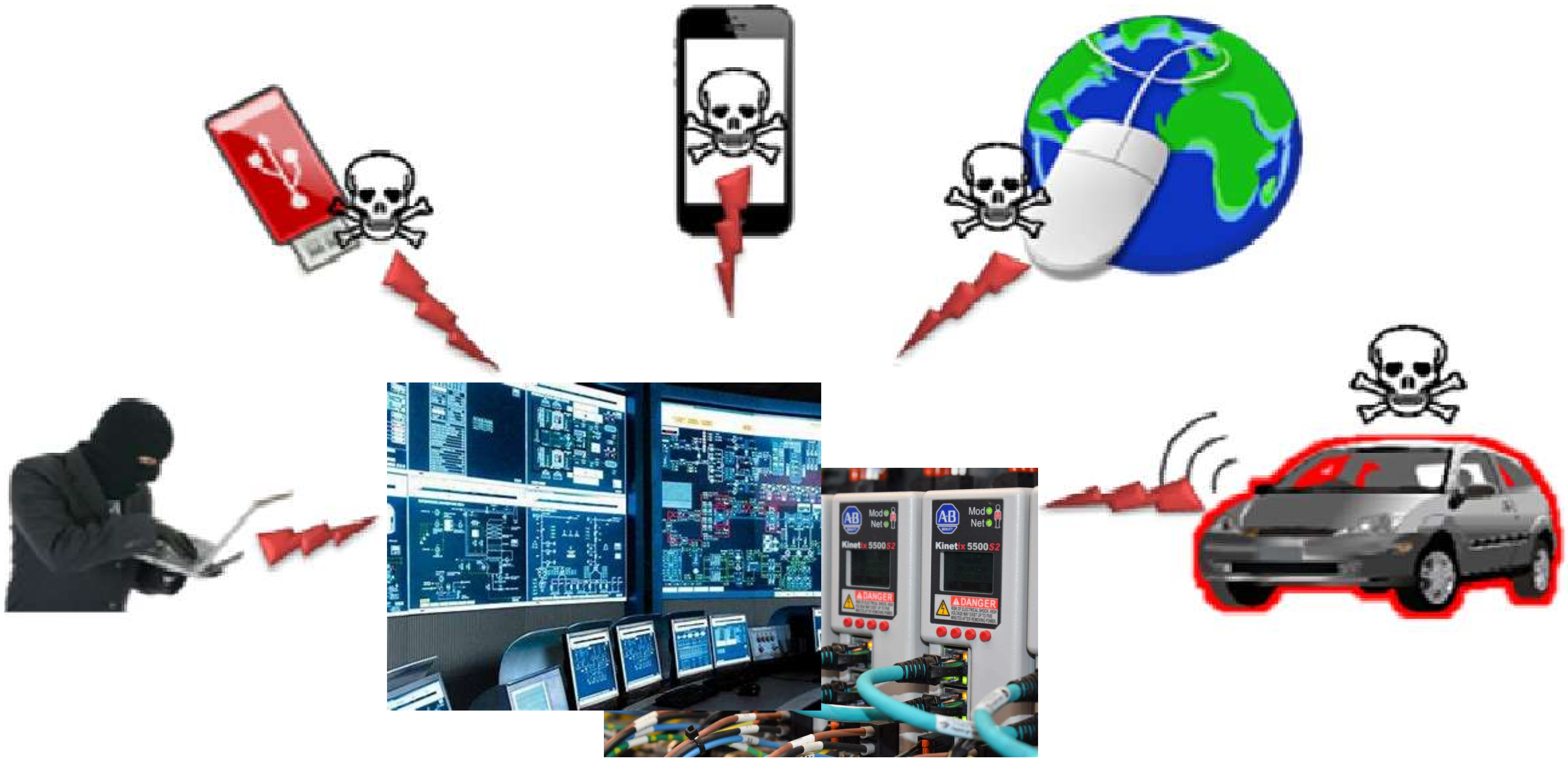


Towards Safe and Secure Autonomous and Cooperative Vehicle Ecosystems. Lima, A; Rocha, F; Volp, M; Verissimo, P. in Proc's 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy (2016, October) @CCS, Vienna-Austria

Overall, are *automated control* ecosystems secure and/or safe?

Or are there relevant gaps?

Safety gap in automated control ecosystems



King Abdullah University of Science and Technology
جامعة الملك عبدالله للعلوم والتقنية

The *SAFETY GAP* in the autonomous vehicles area ...

Safety gap in vehicle ecosystems

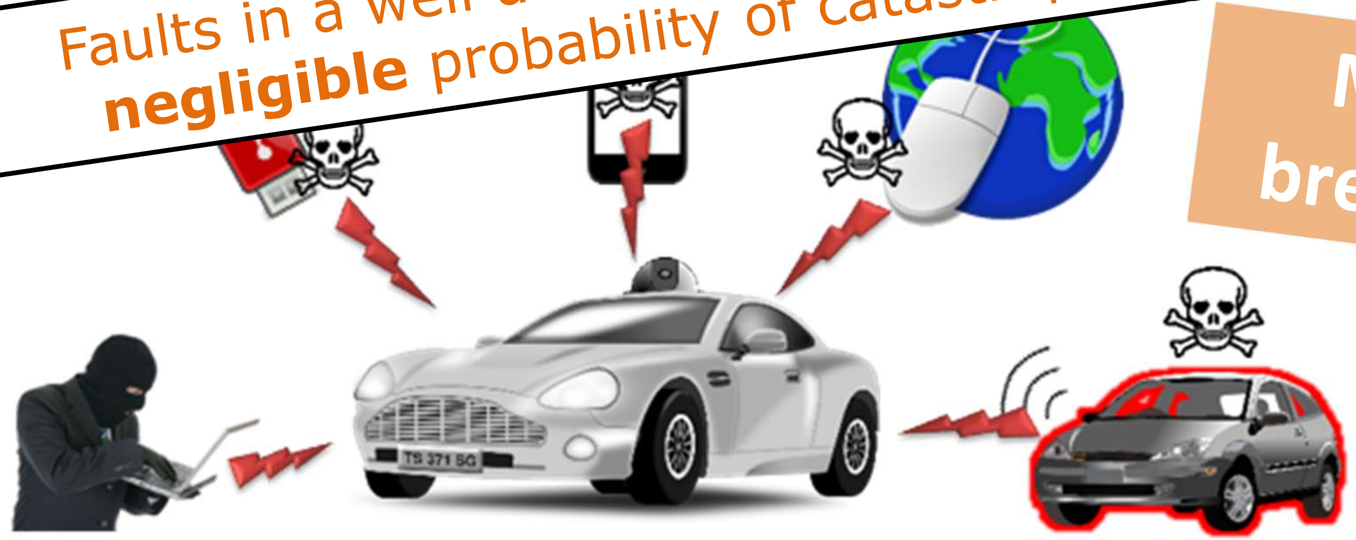
Or...
maybe those reported accidents ... were not really just bad luck?

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

Safety gap in vehicle ecosystems

Faults in a well designed car ecosystem lead to an **infinitesimal and acceptable** probability of catastrophic failure;

Faults in a well designed car may imply a **non-negligible** probability of catastrophic failure



Move fast break things?

Towards Safe and Secure Autonomous and Cooperative Vehicle Ecosystems. Lima, A; Rocha, F; Volp, M; Verissimo, P. in Proc's 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy (2016, October) @CCS, Vienna-Austria

But it can get worse:

The
SAFETY-SECURITY GAP
in the autonomous vehicles area ...
... *(land, air, space)*

Safety-security gap in vehicle ecosystems

Faults in a well designed car ecosystem lead to an **infinitesimal and acceptable** probability of catastrophic failure;

Faults in a well designed car may imply a **non-negligible** probability of catastrophic failure

Vulnerabilities in a car ecosystem **will** lead, rather sooner than later, to catastrophic failures;

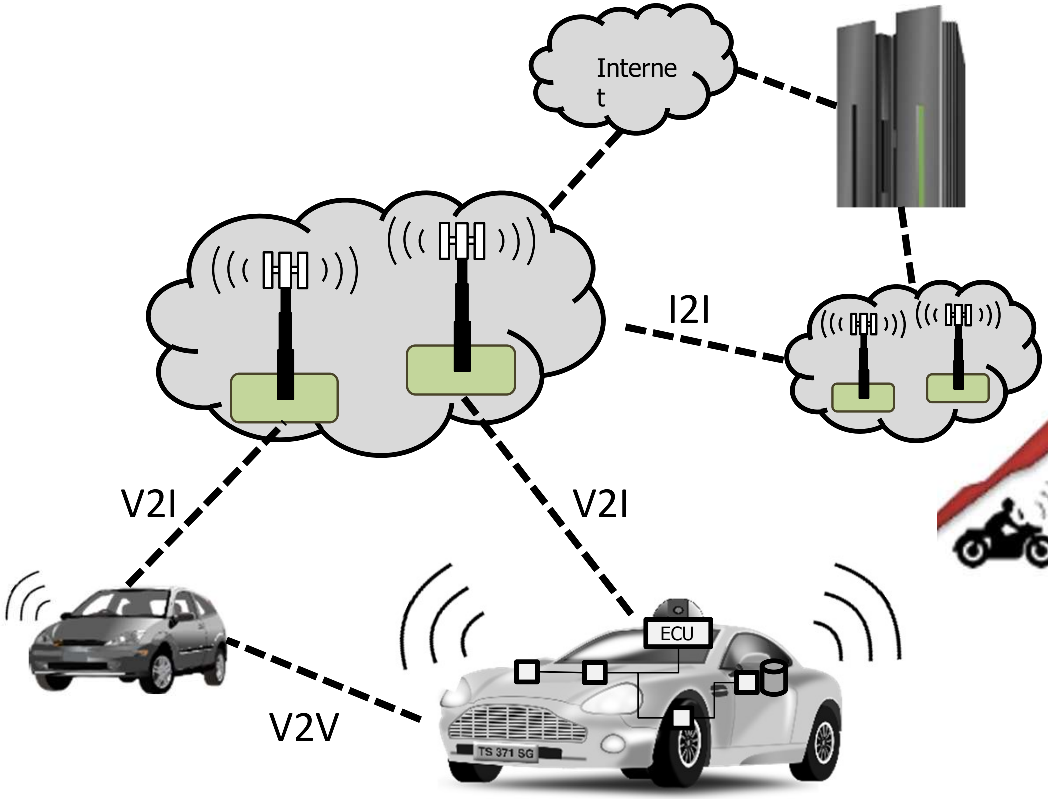
Towards Safe and Secure Autonomous and Cooperative Vehicle Ecosystems. Lima, A; Rocha, F; Volp, M; Verissimo, P. in Proc's 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy (2016, October) @CCS, Vienna-Austria

What is the safety and security ***THREAT SURFACE*** in the autonomous vehicles ***ECOSYSTEM*** ...?

Autonomous Vehicle Ecosystem

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

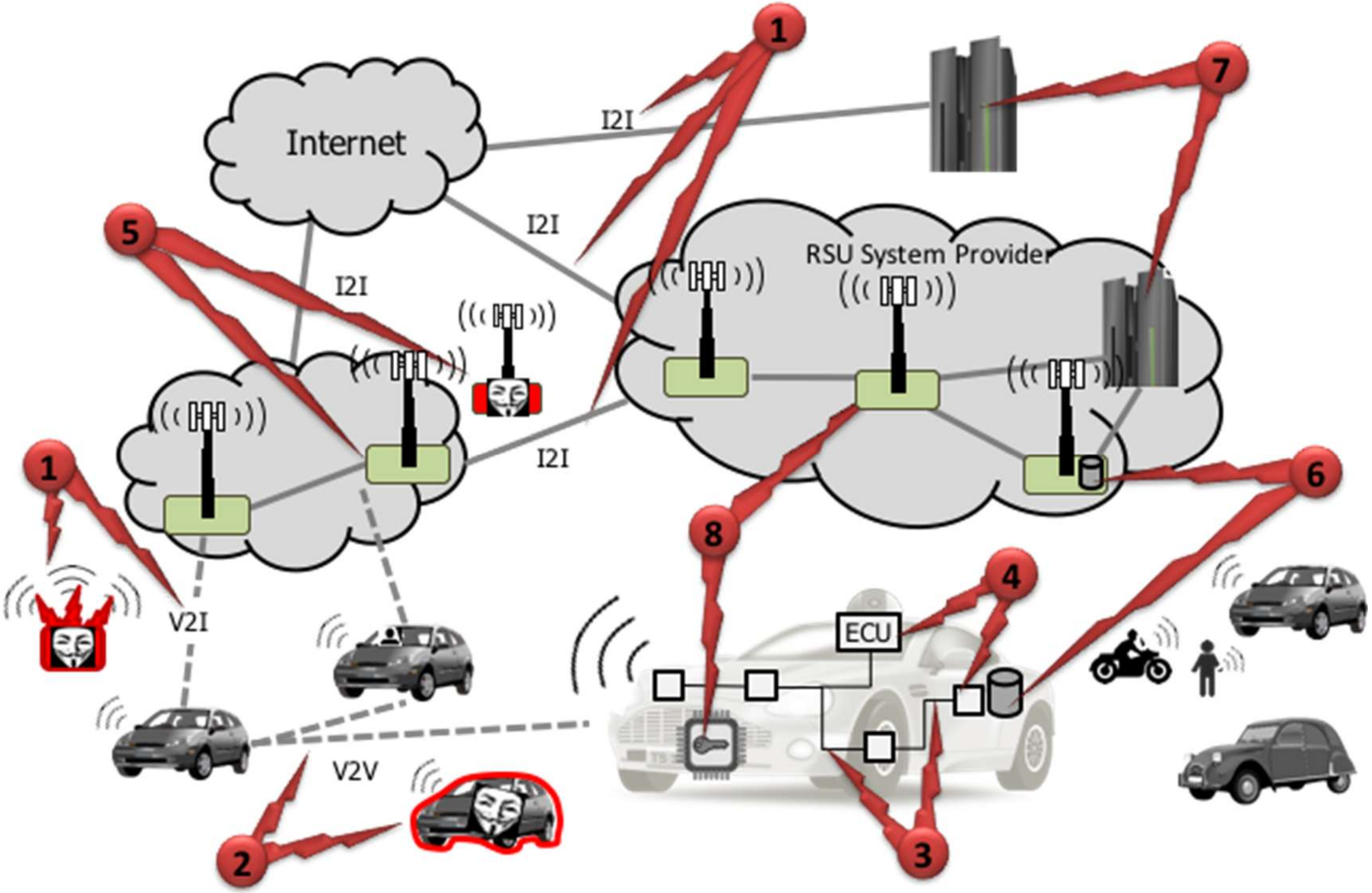
2016



FIRST COMPREHENSIVE STUDY OF THE THREAT SURFACE AND SAFETY-SECURITY GAP OF AUTONOMOUS AND COOPERATIVE VEHICLE ECOSYSTEMS

Towards Safe and Secure Autonomous and Cooperative Vehicle Ecosystems. Lima, A; Rocha, F; Volp, M; Verissimo, P. in Proc's 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy (2016, October) @CCS, Vienna-Austria

Autonomous vehicle ecosystem
threat surface perhaps wider than many think



Threat Vectors

How serious is that?

«IF IT AIN'T SECURE, IT AIN'T SAFE»

Safety-security gap in vehicle ecosystems

Faults in a well designed car ecosystem lead to an **infinitesimal and acceptable** probability of catastrophic failure;

Faults in a well designed car may imply a **non-negligible** probability of catastrophic failure

Vulnerabilities in a car ecosystem **will** lead, rather sooner than later, to catastrophic failures;

Towards Safe and Secure Autonomous and Cooperative Vehicle Ecosystems. Lima, A; Rocha, F; Volp, M; Verissimo, P. in Proc's 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy (2016, October) @CCS, Vienna-Austria

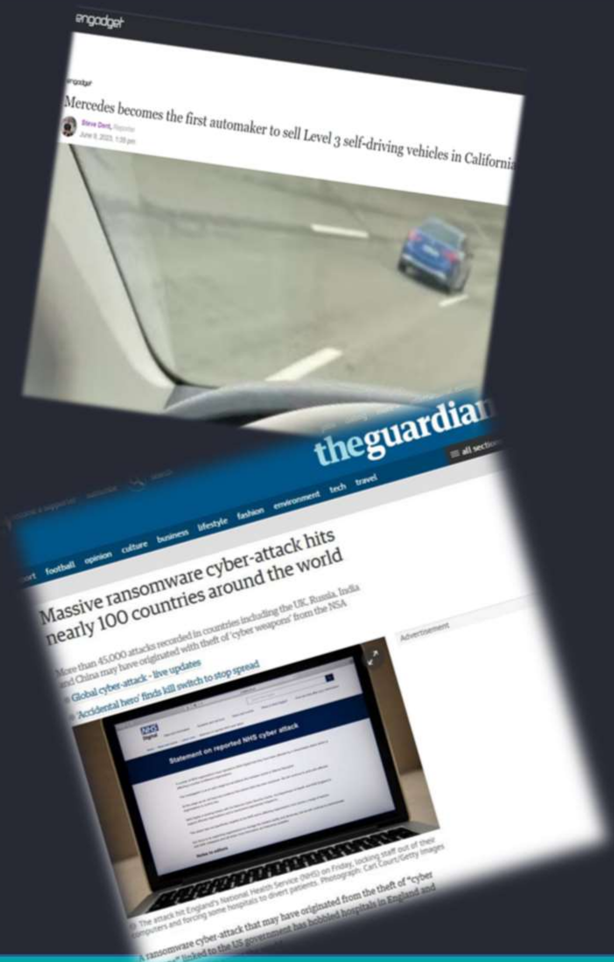
AI as band-aid?

The specific pitfalls of AI/ML for critical systems …

King Abdullah University of Science and Technology | Resilient Computing and Cybersecurity Center

Resilient Computing and Cybersecurity Center

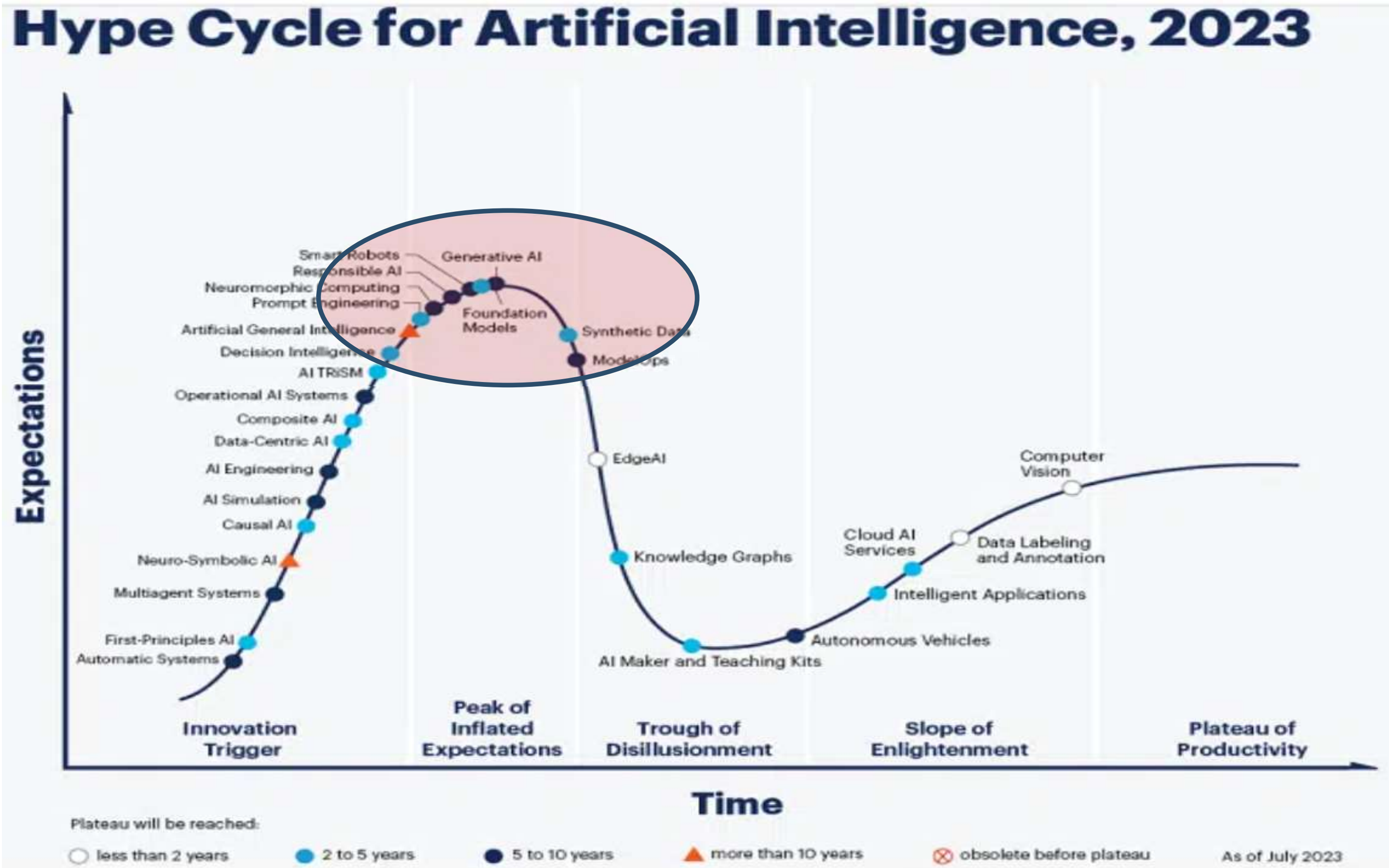
AI/ML vs. Security vs. Safety



AI/ML for Cybersecurity and Safety

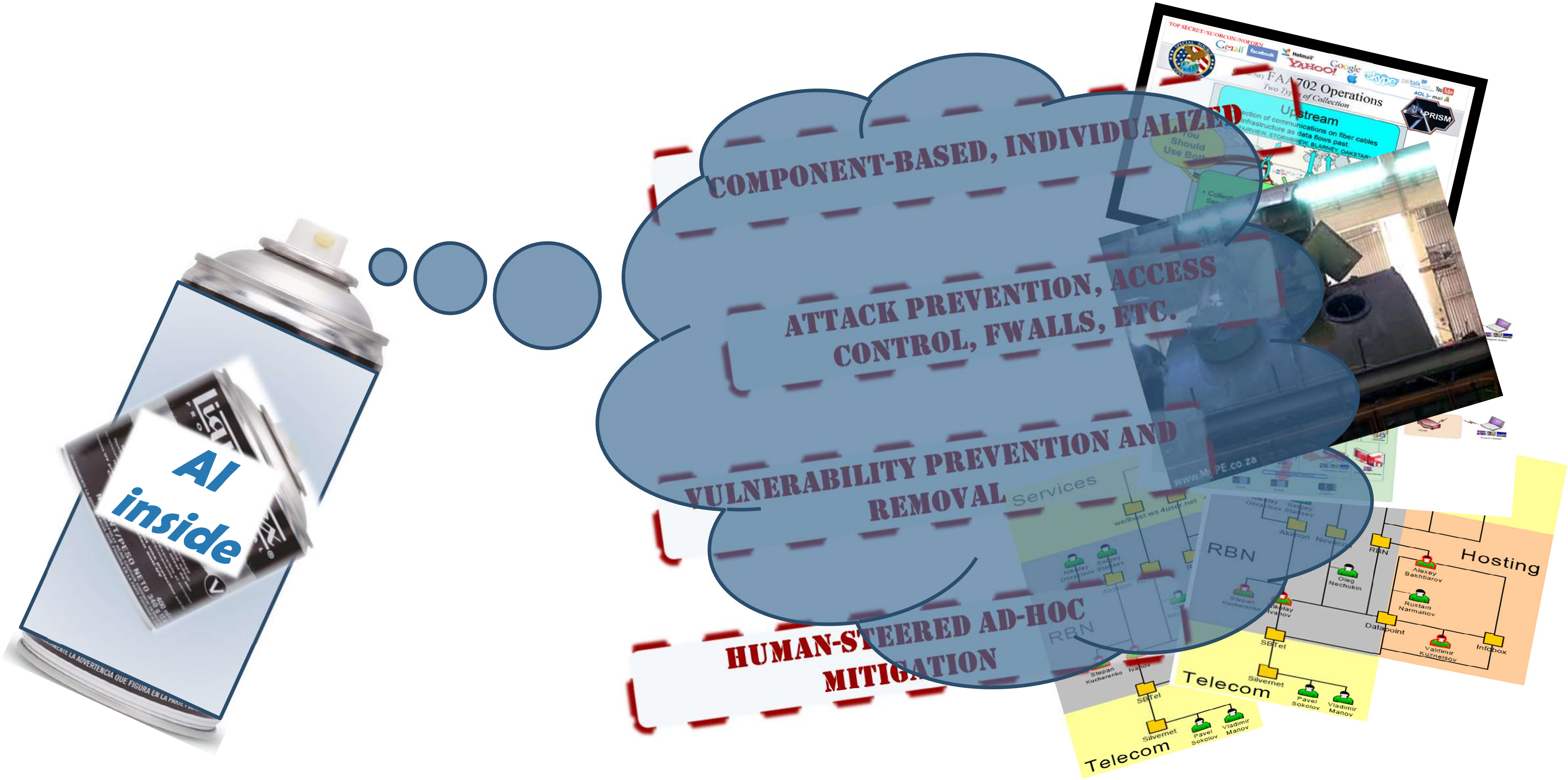
Cybersecure and Safe AI/ML

Enter AI, ML -- Episode II



AI, ML, DNN, LLM, GPT, … to the rescue !!

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology



Some myths and misconceptions about safety and security of autonomous vehicle control systems

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

Some misconceptions about ML-driven AV, on safety or security

- AVs are safer than human-driven vehicles, because AVs don’t do human-like errors
- Image recognition and models pre-trained to all possibly know events are all that’s needed
- Commercial AVs drove over 10Mio Kms, so have actually reached a very good confidence about robustness of their control models w.r.t. Safety

- See examples given...
- Stateless, fragile to unanticipated responses/emergent behavior, open environments unpredictability, semantic & coverage gap (V&V prob)
- To meet 95% safety confidence, 200M miles/fatality: need to test for 600M to 2B miles without seeing a fatality.
- «And you’ll always have camels...»

Some misconceptions about ML-driven AV, on safety or security

- NeuroSymbolic, PhysicsInformed approaches will fix things
- NS and PI improve, but are fixes at data level. In a control system, *system awareness* is paramount.

- Invidualistic cars OK, no need for ecosystem
- Invidualistic cars worsen safety, cooperation is key for AV driving safety

- Security can be fixed as in IT systems
- Without security there is no safety
- Worse in CPS/IoT scenario

Homogeneous ML-based systems cannot give strong assurance and resilience guarantees

King Abdullah University of Science and Technology

- **Status-quo**
  - *Autonomous cars use ML-powered multi-sensor perception (mainly vision) and control, and sometimes redundant modules to which the MLearned module hands over in case of problems.*
- ***Assurance***
  - *LOW- Infeasible to provide reliable figures/conclusions, impossible to certify*
- ***Resilience***
  - *LOW- Fair success in handling unforeseen, emergent or out-of-envelope behaviours; often even blind to those situations*

Tesla vision did not recognize a camel, causing an accident in the UAE

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

LESSONS LEARNT?

Cruise driverless car runs over woman and stops

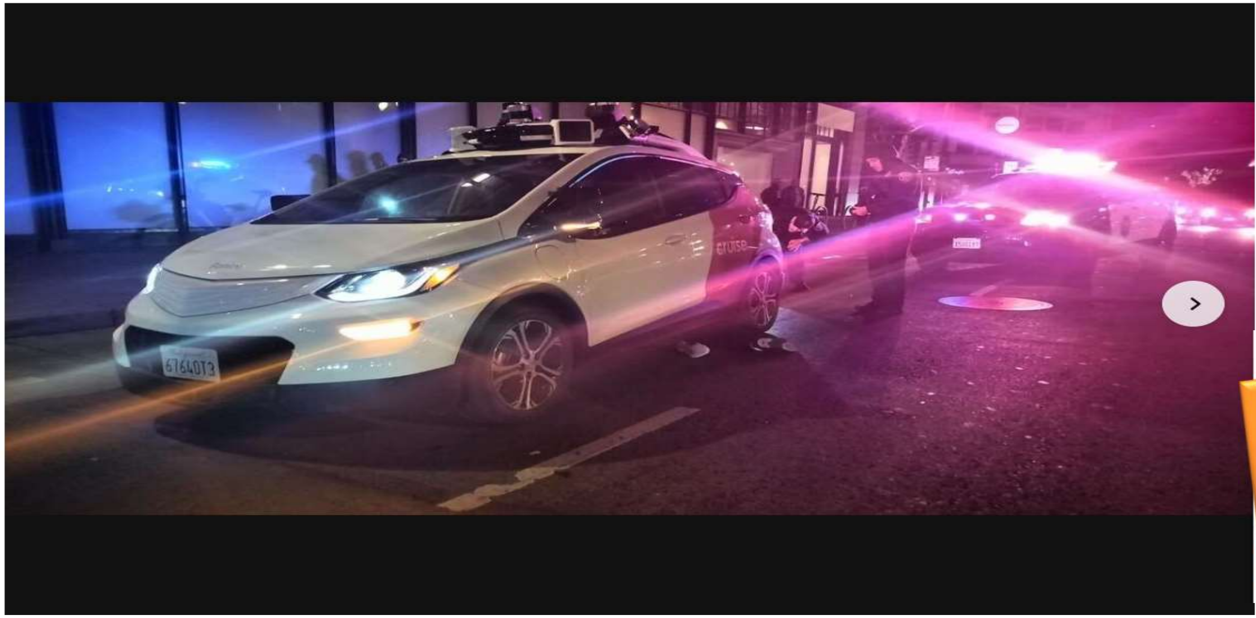
King Abdullah University of Science and Technology

BAY AREA // SAN FRANCISCO

Driver hits woman in S.F., then Cruise driverless car runs her over; photo shows victim trapped

Jordan Parker, Nora Mishanec

Oct. 2, 2023 | Updated: Oct. 3, 2023 3:52 p.m.



LESSONS LEARNT?

One of Uber's Self-Driving Cars Hit and Killed a Woman in Arizona

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

Self-driving Uber car that hit and killed woman did not recognize that pedestrians jaywalk

The automated car lacked "the capability to classify an object as a pedestrian unless that object was near a crosswalk," an NTSB report said.

DISASTERS

Dashcam video of deadly self-driving Uber crash released

By Nicole Darrah , · Fox News

Published March 22, 2018 12:45am EDT | Updated March 22, 2018 12:59am EDT

Dashcam catches the moment self-driving Uber hits pedestrian

LESSONS LEARNT?

The serious ecosystem security risks

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

LESSONS LEARNT?

Philosophical side of the problem:

«Control the physics of event interleaving in autonomous object ecosystems, acting in real time, in open and largely unpredictable environments»

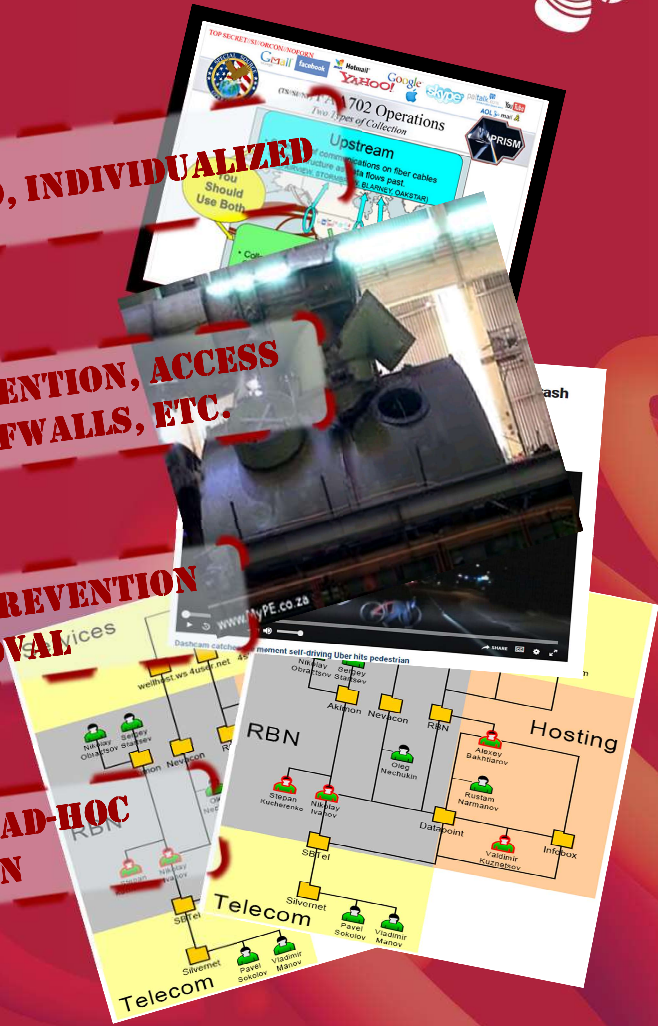
Solutions? ...

COMPONENT-BASED, INDIVIDUALIZED

ATTACK PREVENTION, ACCESS CONTROL, FWALLS, ETC.

VULNERABILITY PREVENTION AND REMOVAL

HUMAN-STEERED AD-HOC MITIGATION



A part of the long journey towards

RESILIENT AUTONOMOUS VEHICLE ECOSYSTEMS

More recently, A. Shoker and R. Yasmin at CybeResil@KAUST, M.Voelp CRITIX@UNILU, V. Rahli @U.BIRMINGHAM, J. Decouchant@U.DELFT

CORTEX Project Info

[2001-04]

INFORMATION SOCIETY TECHNOLOGIES (IST) PROGRAMME

Project acronym: ***CORTEX***
Project full title:
CO-operating Real-time senTient objects: architecture and EXperimental evaluation

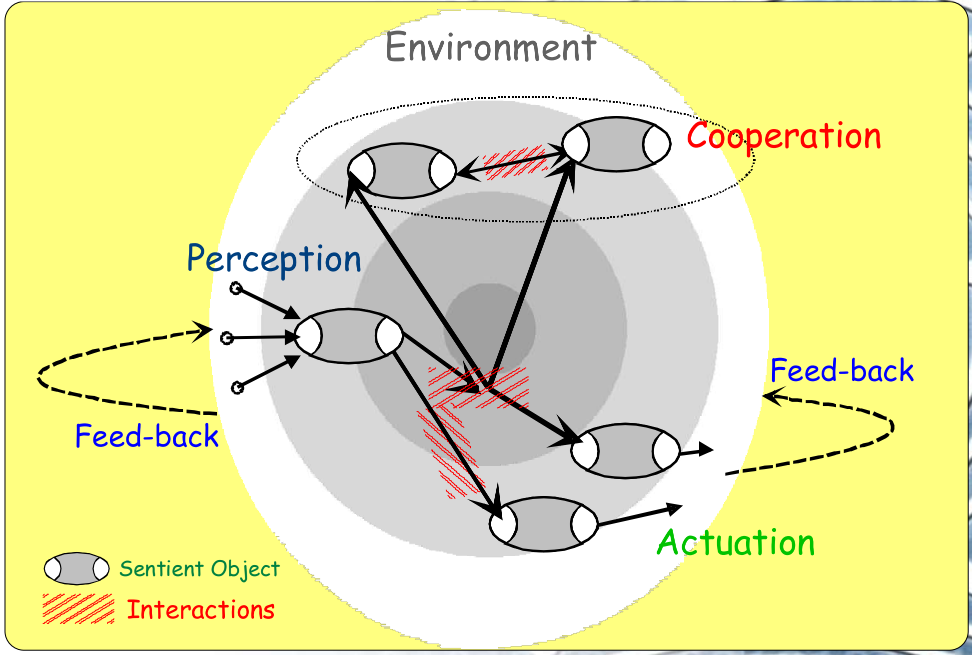
- Members:
  - Univ. Lisboa Fac. Of Sciences (PT) **(proj. coord.)**
  - Trinity College of Dublin (IR)
  - U. of Lancaster (UK)
  - U. of Ulm (DE)
- Duration:
  - 3 years, starting April 2001
- Budget:
  - 2 MEURO

'Sentient objects' interaction model

Abstract safe distributed real-time (DRT) autonomous control of free-running objects

should support the classes of R/T interactions objects need to perform:

- sentience of body and of environment;
- environment-to-object and vice-versa;
- object-to-object



[P. Veríssimo and A. Casimiro. The Timely Computing Base Model and Architecture. IEEE Tacs. on Computers, 2002]

Overarching predicates

Generic predicates dictate system correctness in face of uncertainty, regardless of functional semantics

No-Contamination - violation of *normal* properties can happen (e.g. timeliness) but never entails violation of *critical* properties (e.g. logical safety)

Property No-Contamination. *Given a history $\mathcal{H}(\mathcal{T}_{\mathcal{P}})$ derived from property $\mathcal{P} \in \mathcal{P}_A$, $\mathcal{H}$ has no-contamination iff, for any timing failure in any execution $X \in \mathcal{H}$, no safety property in $\mathcal{P}_A$ is violated.*

Coverage Stability – the coverage *(less than or equal to one)* of any property (e.g. timeliness) remains stable within bounds

Property Coverage Stability. *Given a history $\mathcal{H}(\mathcal{T}_{\mathcal{P}})$ derived from property $\mathcal{P} \in \mathcal{P}_A$, with assumed coverage $P_{\mathcal{P}}$, $\mathcal{H}$ has coverage stability iff the set of executions contained in $\mathcal{H}$ is timely with a probability $p_{\mathcal{H}}$ such that $|p_{\mathcal{H}} - P_{\mathcal{P}}| \leq p_{dev}$, for p_{dev} known and bounded.*

Dependable adaptation at work : Some fairly complete behaviour classes

- Define behaviour classes with regard to a property P:
- Adaptive
  - Recurrent violation of property P is accepted, if with a known and bounded degree and/or probability

[Reconfigur. and adapt., Casimiro et al., SRDS'01]

- Safe
  - Occasional violation of property P is accepted, if the system can react dependably

[Timing error masking , Casimiro et al., DSN'02]

- Fail-safe
  - Any violation of property P is not acceptable and so the system must do a fail-safe/op routine (e.g. stop)

[Fail-safe operation, Casimiro et al., DSN'00]

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

KARYON PROJECT: Kernel-Based ARchitecture for safetY-critical cONtrol

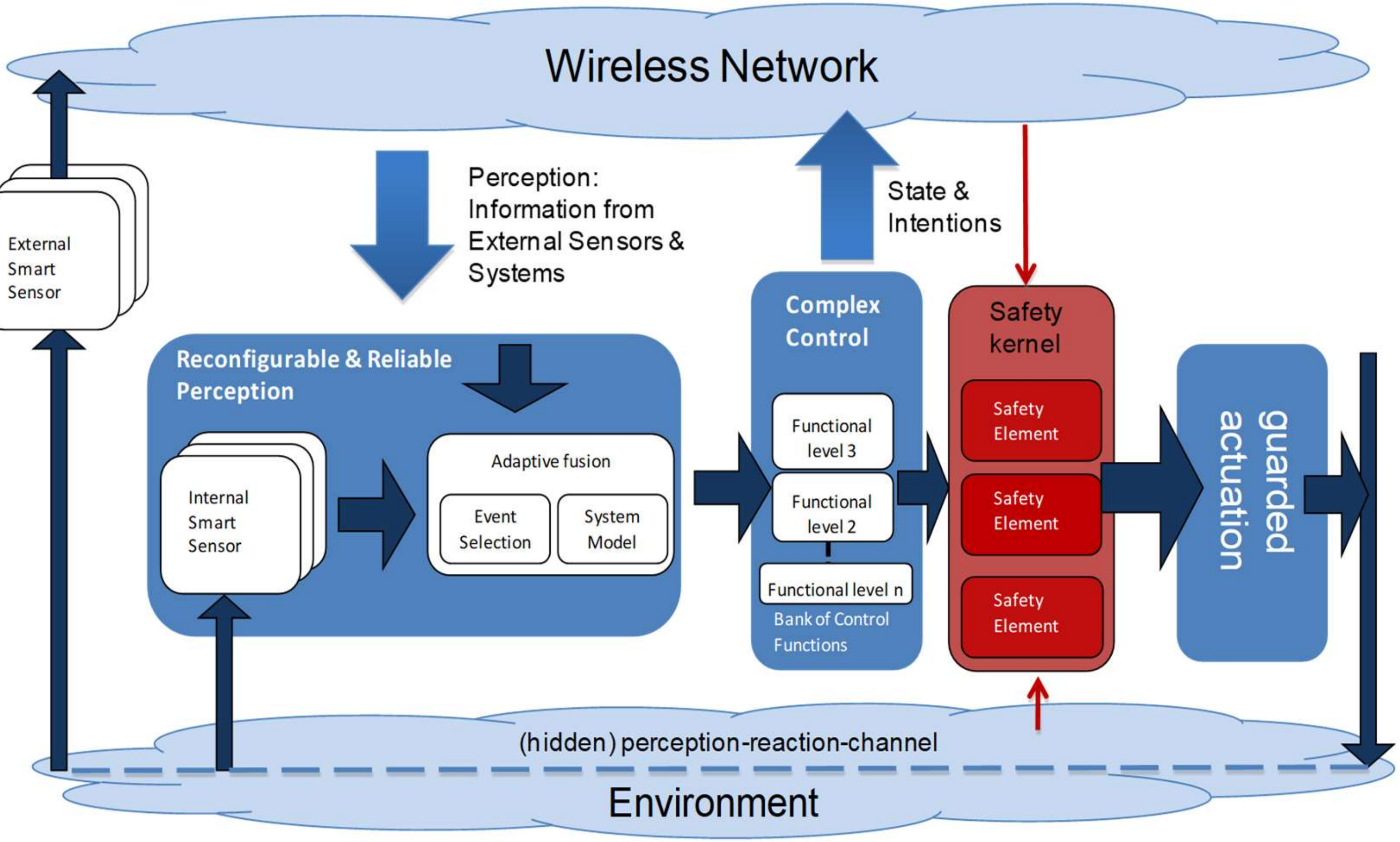
2011-2014

Academia & Research Institutes SMEs and Industry	Proof-of-concept prototypes Simulations
FACULDADE DE CIÊNCIAS UNIVERSIDADE DE LISBOA; gmv INNOVATING SOLUTIONS; OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG; EMBRAER	**Avionics** UAS/Aircraft flight mission
SP your Science Partner; CHALMERS UNIVERSITY OF TECHNOLOGY; 4S Group Technology for Sustainability	**Automotive** Adaptive cruise control Coordinated lane change Coordinated intersection crossing

- Provide system solutions for predictable and safe coordination of smart vehicles that autonomously cooperate and interact in an open and inherently uncertain environment

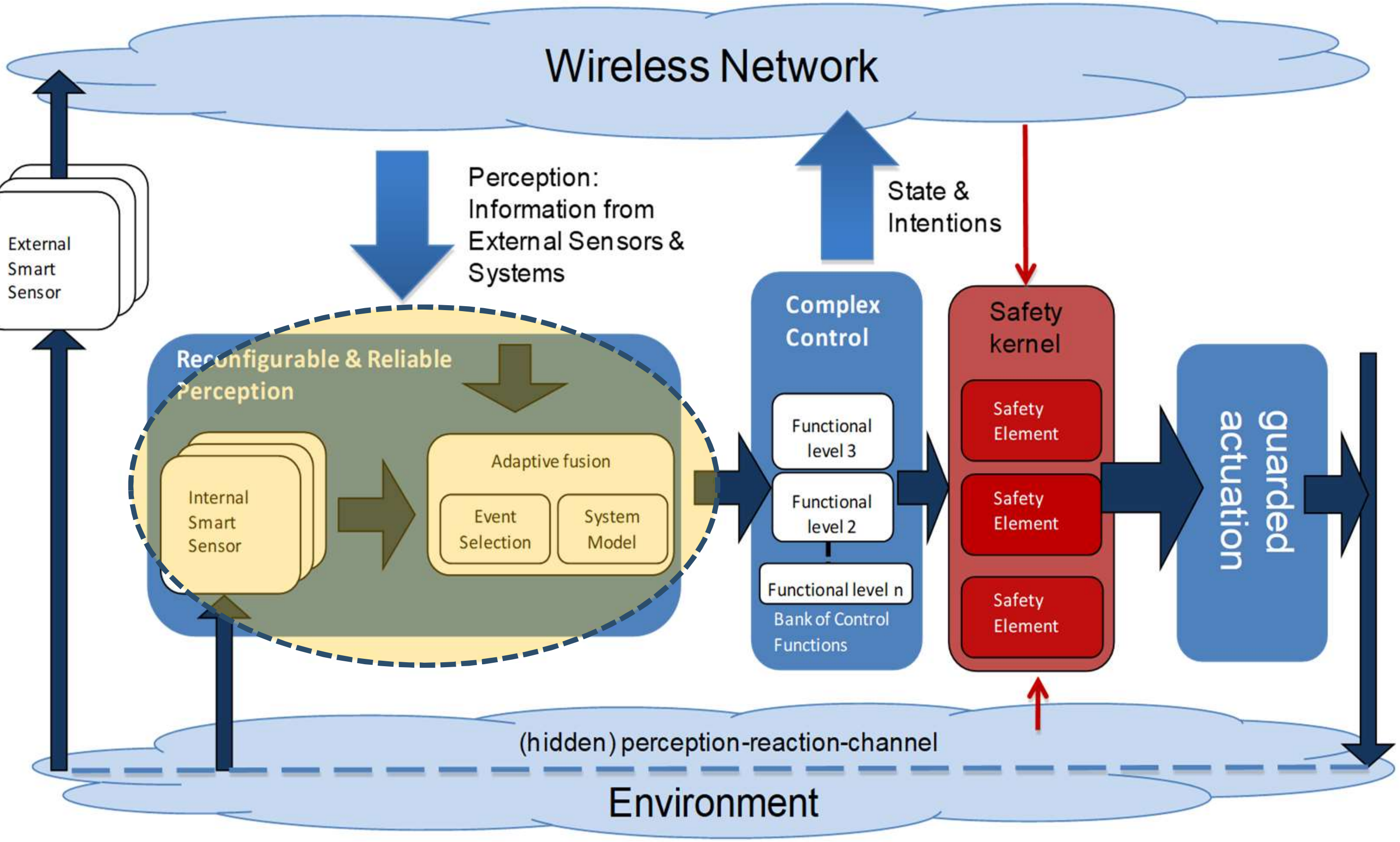
93

KARYON architectural view: proof of concept of hybridisation for safety



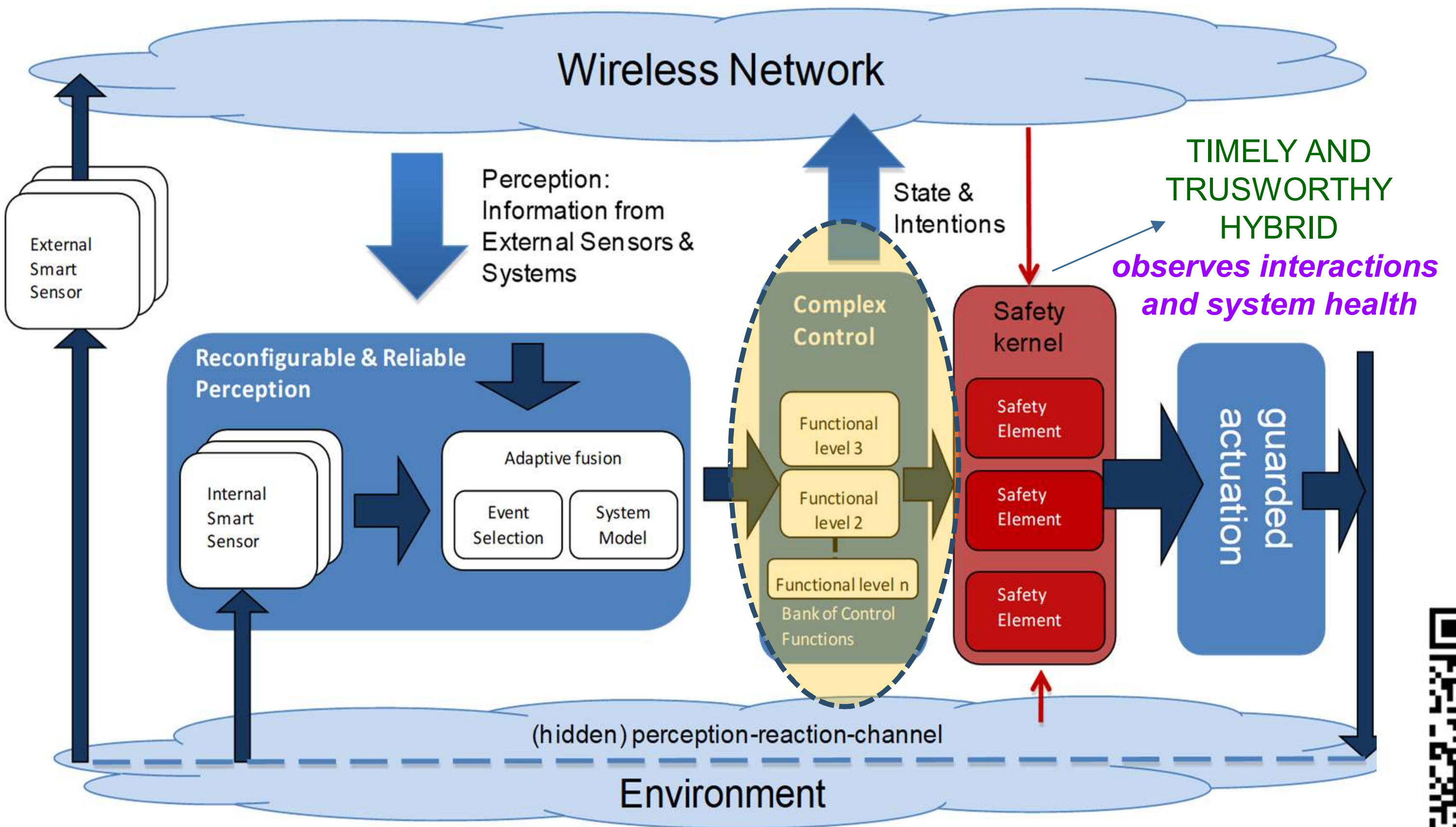
*A. Casimiro, J. Kaiser, E. Schiller, P. Costa, J. Parizi, R. Johansson, R. Librino, "**The KARYON Project: Predictable and Safe Coordination in Cooperative Vehicular Systems**", in 2nd Workshop on Open Resilient Human-aware CPS (WORCS'13), Jun. 2013.*

KARYON architectural view: proof of concept of hybridisation for safety



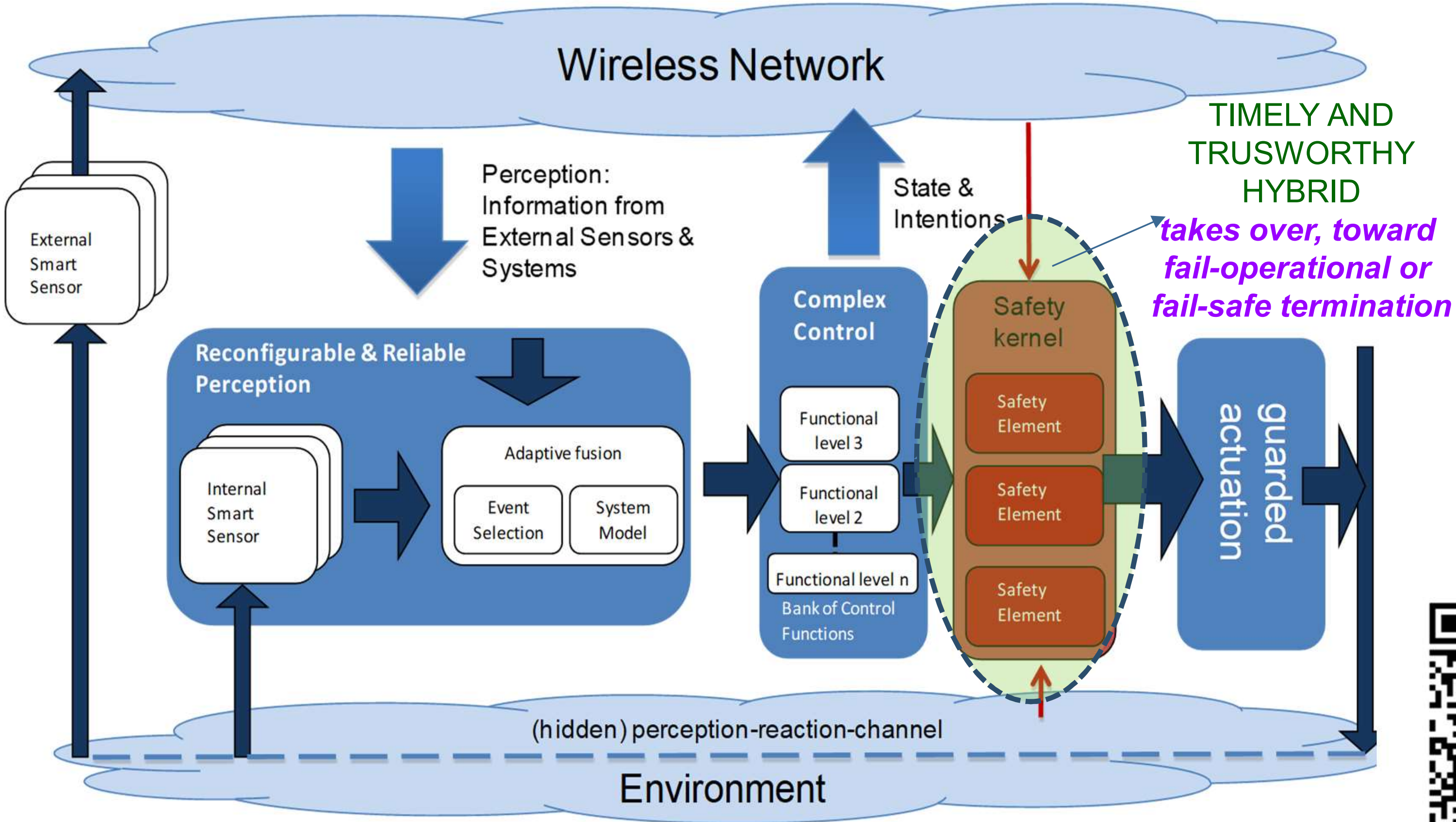
*A. Casimiro, J. Kaiser, E. Schiller, P. Costa, J. Parizi, R. Johansson, R. Librino, "**The KARYON Project: Predictable and Safe Coordination in Cooperative Vehicular Systems**", in 2nd Workshop on Open Resilient Human-aware CPS (WORCS'13), Jun. 2013.*

KARYON architectural view: proof of concept of hybridisation for safety



A. Casimiro, J. Kaiser, E. Schiller, P. Costa, J. Parizi, R. Johansson, R. Librino, "**The KARYON Project: Predictable and Safe Coordination in Cooperative Vehicular Systems**", in 2nd Workshop on Open Resilient Human-aware CPS (WORCS'13), Jun. 2013.

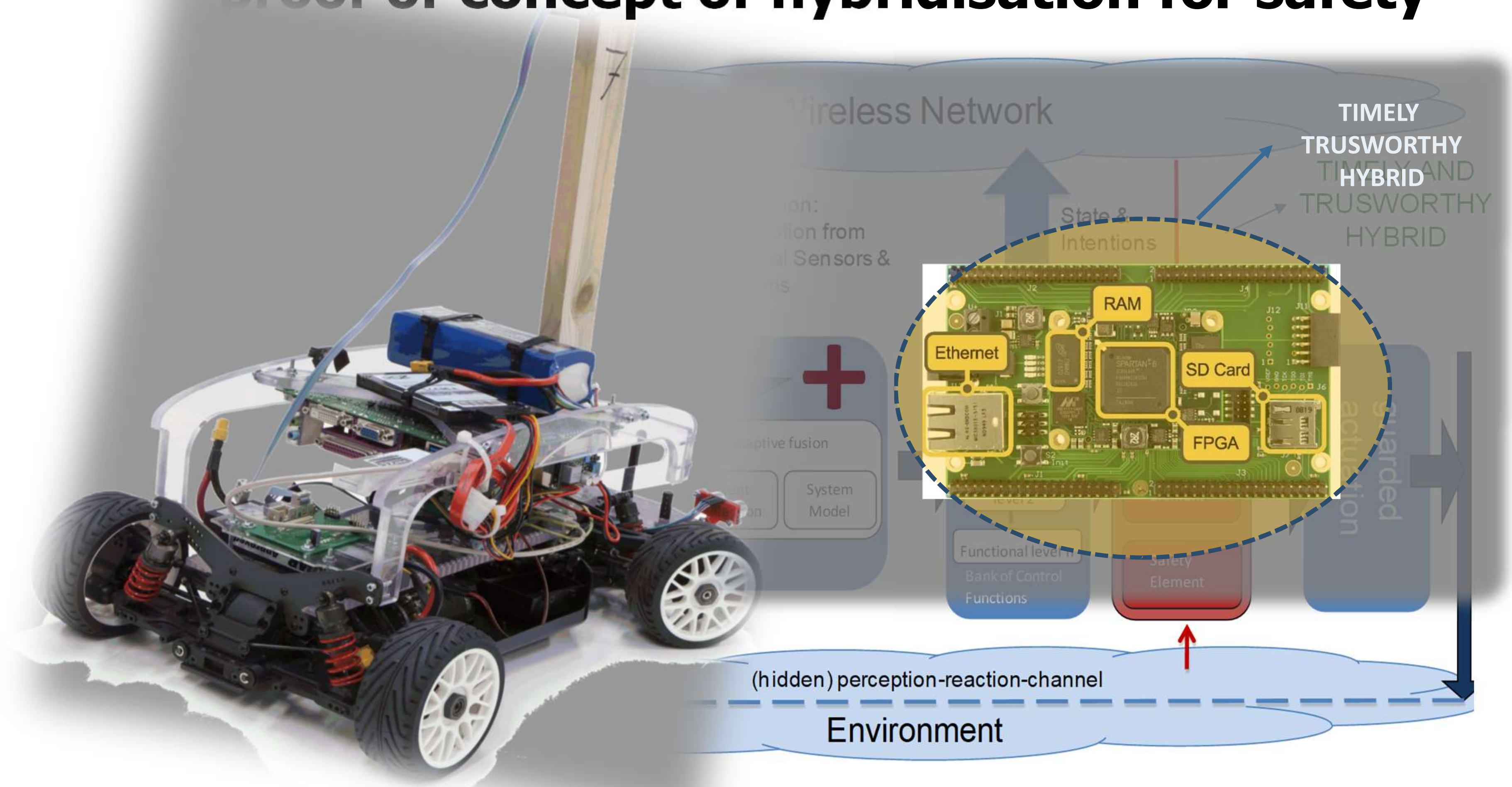
KARYON architectural view: proof of concept of hybridisation for safety



*A. Casimiro, J. Kaiser, E. Schiller, P. Costa, J. Parizi, R. Johansson, R. Librino, "**The KARYON Project: Predictable and Safe Coordination in Cooperative Vehicular Systems**", in 2nd Workshop on Open Resilient Human-aware CPS (WORCS'13), Jun. 2013.*

KARYON architectural view: proof of concept of hybridisation for safety

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology



A. Casimiro, J. Kaiser, E. Schiller, P. Costa, J. Parizi, R. Johansson, R. Librino, "**The KARYON Project: Predictable and Safe Coordination in Cooperative Vehicular Systems**", *in 2nd Workshop on Open Resilient Human-aware CPS (WORCS'13), Jun. 2013.*

Intel Collaborative Research Institute for

Collaborative Autonomous & Resilient Systems *(CARS)*

securityandtrust.lu

CRITIX

2017-2020

https://www.icri-cars.org/

ICRI-CARS » Resilient Autonomy » Mission

ICRI-CARS

- Mission
- Research Topics
- Principal Investigators
- TU Darmstadt
- Aalto University
- Ruhr-University Bochum
- Critix@ University of Luxembourg
- TU Wien
- Collaborations

Intel Collaborative Research Institute for Collaborative Autonomous & Resilient Systems (ICRI-CARS)

About Collaborative Autonomous and Resilient Systems (CARS)

The mission of the ICRI-CARS is the study of security, privacy, and safety of autonomous systems that may collaborate with each other. Examples include drones, self-driving vehicles, or collaborative systems in industrial automation. CARS introduce a new paradigm to computing that is different from conventional systems in a very important way: they must learn, adapt, and evolve with minimal or no supervision. A fundamental question therefore, is what rules and principles should guide the evolution of CARS?

This raises security related questions in multiple research areas:

1. Trustworthy and Controllable Autonomy
2. Fair and Safe Collaboration Tolerating Failures and Attacks
3. Intelligent Security Strategies for Self-Defense and Self-Repair
4. Integration of Safety, Security, and Real-time Guarantees
5. Autonomous Systems, Ecosystem Scenarios, Requirements, Case Studies, and Validation
6. Advanced Platform Security for Long-term Autonomy

Resilience enablers for autonomous and collaborative vehicles

Applied safe and secure DRT autonomous control --- general driving

- **Powerful architectures** (e.g. manycores), capable of: high-power computing, enabling security/safety defenses
- **Secure and dependable *real-time* communication**, V2V and V2I, despite accidents and attacks
- **Automatic in-car resilience** mechanisms for safety and security (gateway, ECU, trusted components/enclaves)

Resilience enablers for autonomous and collaborative vehicles

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

- **Powerful architectures** (e.g. manycores), capable of: high-power computing, enabling security/safety defenses
- **Secure and dependable *real-time* communication**, V2V and V2I, despite accidents and attacks
- **Automatic in-car resilience** mechanisms for safety and security (gateway, ECU, trusted components/enclaves)

Resilience enablers for autonomous and collaborative vehicles

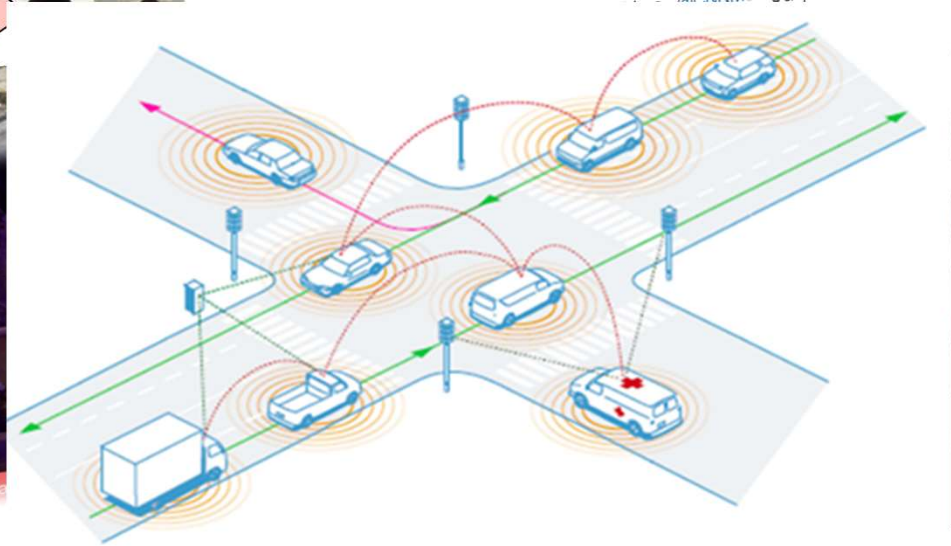
جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

- **Powerful architectures** (e.g. manycores), capable of: high-power computing, enabling security/safety defenses
- **Secure and dependable *real-time* communication**, V2V and V2I, despite accidents and attacks
- **Automatic in-car resilience** mechanisms for safety and security (gateway, ECU, trusted components/enclaves)

Ecosystem approach: Cooperation is key!

Individualistic cars worsen safety!

Cooperation is key!



Towards Safe and Secure Autonomous and Cooperative Vehicle Ecosystems. Lima, A; Rocha, F; Volp, M; Verissimo, P. in Proc's 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy (2016, October) @CCS, Vienna-Austria

Real-Time and Byzantine Resilient Digital Twins:
Beyond mere SCADA near-Real-Time Data Dissemination

securityandtrust.lu

CRITIX

Modularly build

PISTIS: Real-Time Byzantine Atomic Broadcast

PISTIS

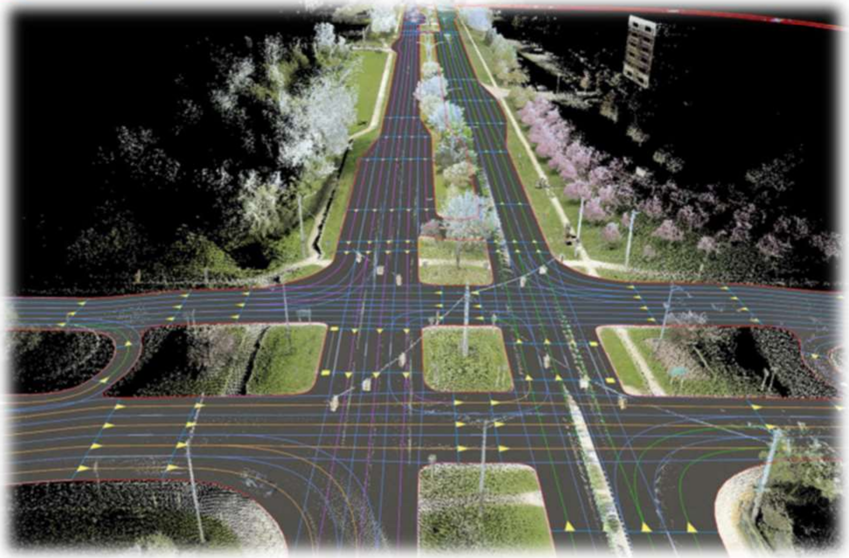
Real-Time Byzantine Consensus

RT-ByzCast: Real-Time Byzantine-Resilient Reliable Broadcast

RT-ByzCast

Poorly behaving network (unbounded probabilistic losses)

intel | Research Institute for Collaborative Autonomous and Resilient Systems



Accurate Real-Time Digital Maps for Autonomous Driving

*D. Kozhaya, J. Decouchant and P. Esteves-Veríssimo, "RT-ByzCast: Real-Time Byzantine-Resilient Reliable Broadcast" , IEEE Transactions on Computers 2019, Core A**

*Kozhaya, D., Decouchant, J., Rahli, V., & Esteves-Verissimo, P. (2021). PISTIS: An Event-Triggered Real-time Byzantine Resilient Protocol Suite. IEEE TPDS. doi:10.1109/tpds.2021.3056718, Core A**

Real-Time and Byzantine Resilient Digital Twins:
Beyond mere SCADA near-Real-Time Data Dissemination

securityandtrust.lu

CRITIX

Modularly build

PISTIS: Real-Time Byzantine Atomic Broadcast

Real-Time Byzantine Consensus

RT-ByzCast: Real-Time Byzantine Resilient Reliable Broadcast

Poorly behaving network (unbounded probabilistic losses)

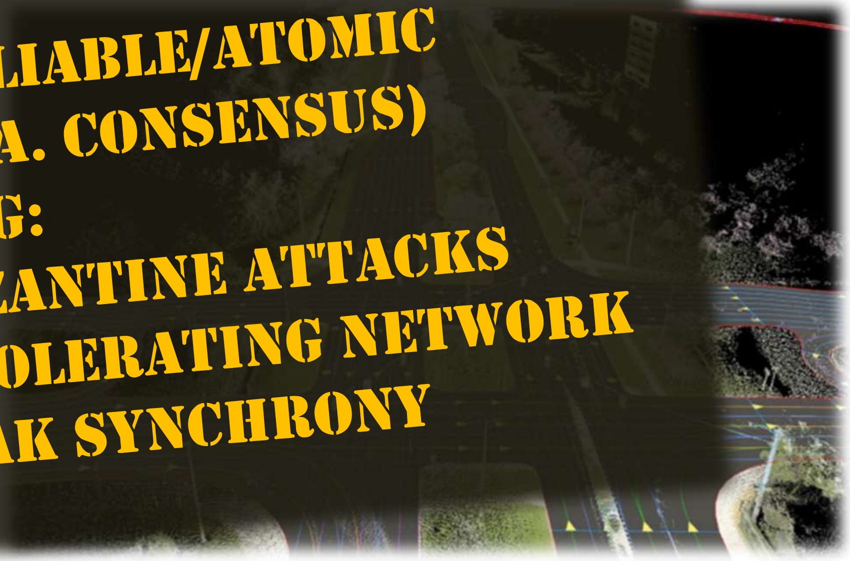
PISTIS

RT-ByzCast

intel | Research Institute for Collaborative Autonomous and Resilient Systems

WORLD-FIRST BYZANTINE RELIABLE/ATOMIC BROADCAST PROTOCOL (A.K.A. CONSENSUS) SIMULTANEOUSLY PROVIDING:

- **RESILIENCE AGAINST BYZANTINE ATTACKS**
- **REAL-TIME OPERATION TOLERATING NETWORK UNCERTAINTIES AND WEAK SYNCHRONY**



Accurate Real-Time Digital Maps for Autonomous Driving

*D. Kozhaya, J. Decouchant and P. Esteves-Veríssimo, "RT-ByzCast: Real-Time Byzantine-Resilient Reliable Broadcast" , IEEE Transactions on Computers 2019, Core A**

*Kozhaya, D., Decouchant, J., Rahli, V., & Esteves-Verissimo, P. (2021). PISTIS: An Event-Triggered Real-time Byzantine Resilient Protocol Suite. IEEE TPDS. doi:10.1109/tpds.2021.3056718, Core A**

Resilience enablers
for autonomous and collaborative vehicles

- **Powerful architectures** (e.g. manycores), capable of: high-power computing, enabling security/safety defenses
- **Secure and dependable *real-time* communication**, V2V and V2I, despite accidents and attacks
- **Automatic in-car resilience** mechanisms for safety and security (gateway, ECU, trusted components/enclaves)

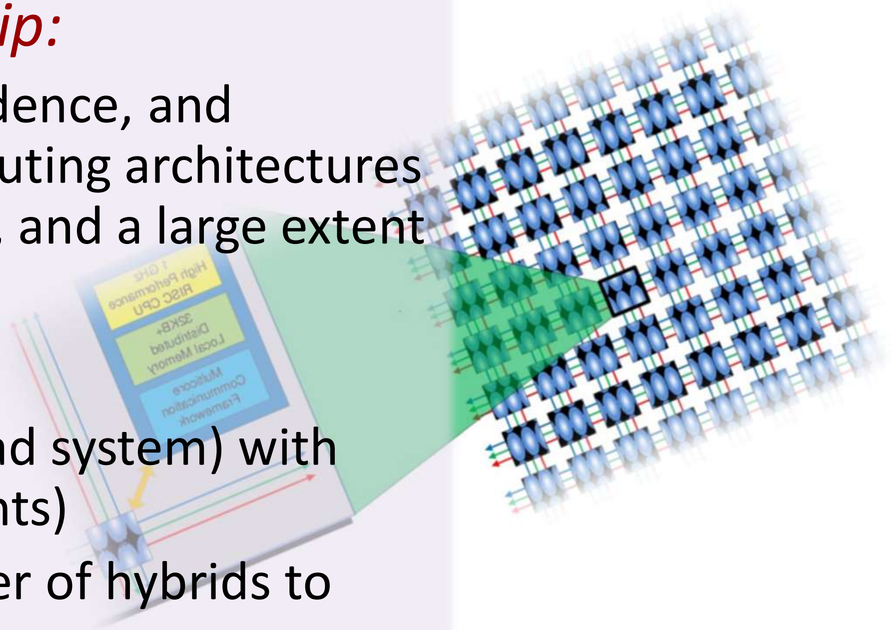
MORE AHEAD!

Resilient, Fault and Intrusion Tolerant Distributed Systems-on-a-Chip (DisSoC)

Manycore Architectures

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

- *Fault-free system designs are infeasible or bearing extreme costs, even if microhypervisor-based*
- *Manycores as distributed-systems-on-a-chip:*
  - Leveraging natural redundancy, fault independence, and diversity, toward extremely dependable computing architectures withstanding advanced and persistent threats, and a large extent of hardware-level faults
- *Hybrid system architecting*
  - Reconcile carefully designed (the larger payload system) with formally verified (the small, trusted components)
  - Hybridisation-aware algorithms leverage power of hybrids to sustain correctness of the whole



Intrusion Resilience System (IRS)

Trustworthy Autonomous Vehicles Architecture (SAVVY)

KAUST In-house Projects

2021----

Towards sustainable security and safety *In AV control*

Towards sustainable security and safety

(inspired by precursor projects Karyon (EU) and ICRI CARS (INTEL)

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

Resilient DRT autonomous control --- general driving

Collaboration among autonomous vehicles (V2V, V2I)

Fault and intrusion tolerant control in-vehicle by eliminating SPOFs, in particular at operating-system level

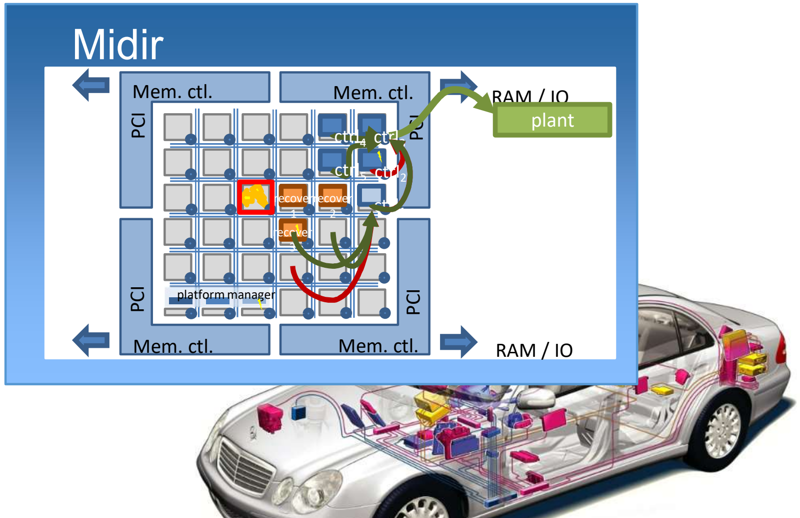


Image credit: Mercedes-Benz Museum (as cited in Computer History Museum, 2011)
Slide from Intel ADG

Intrusion Resilience System (IRS)

The Concept: intrusion masking for real-time fault and intrusion tolerance (R/T FIT)

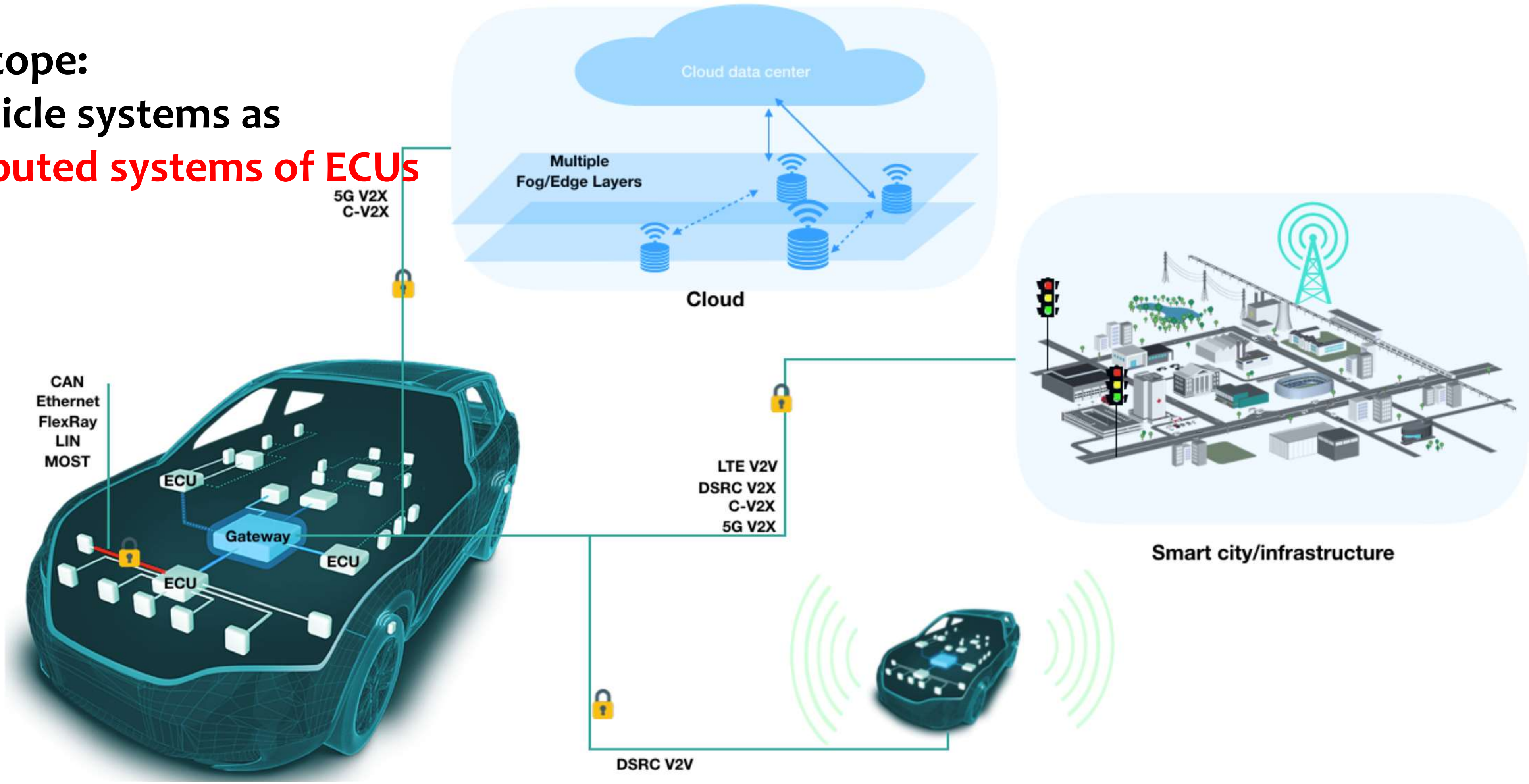
جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

- IRS as a **distributed** service/middleware/library securing critical real-time in-car applications
- ***Distributed State Machines*** over a number of diverse ECUs

A. Shoker, V. Rahli, J. Decouchant and P. Esteves-Verissimo, "Intrusion Resilience Systems for Modern Vehicles," 2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring), Florence, Italy, 2023

Automotive Ecosystem

Our scope:
In-vehicle systems as
Distributed systems of ECUs



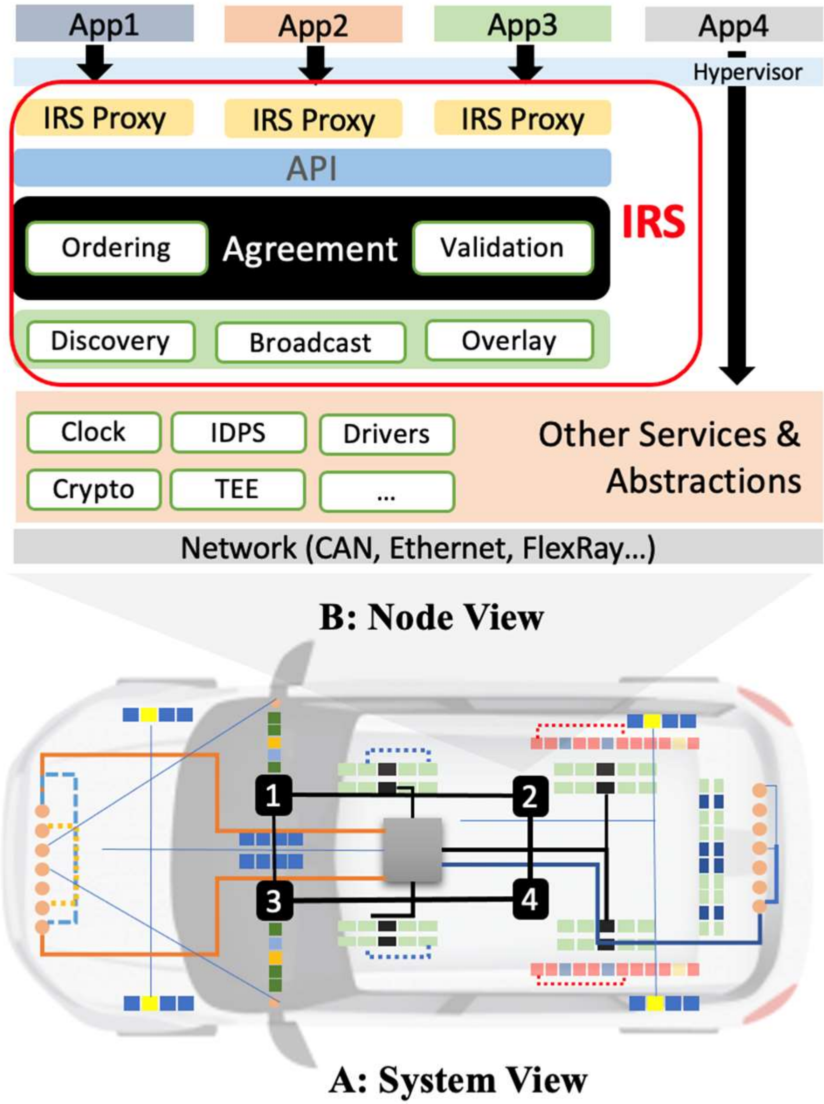
Digitalization, an automation enabler

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of Science and Technology

More than **100**	**SENSORS & ACTUATORS** LIDAR, camera, air, temprature, engine, oil, throttle, spark, valve, lamp, etc.
More than **100**	**ELECTRONIC CONTROLERS** Body, engine, doors, windows, seats, airbag, mirrors, chassis, telecom, voice, mic, etc.
More than **10**	**NETWORKS CONNECTED** CAN, CAN FD, CAN XL, Automotive Ethernet, FlexRay, LIN, MOST,
More than **3 M**	**FUNCTIONS OF CODE (Volvo)** - OS (LynxOS, Neutrino, AGL, Android auto, Apple CarPlay) - Virtualization hypervisors - Applications (ADAS, infotainment, Android, Apple)

Intrusion Resilience System (IRS) The Concept: intrusion masking

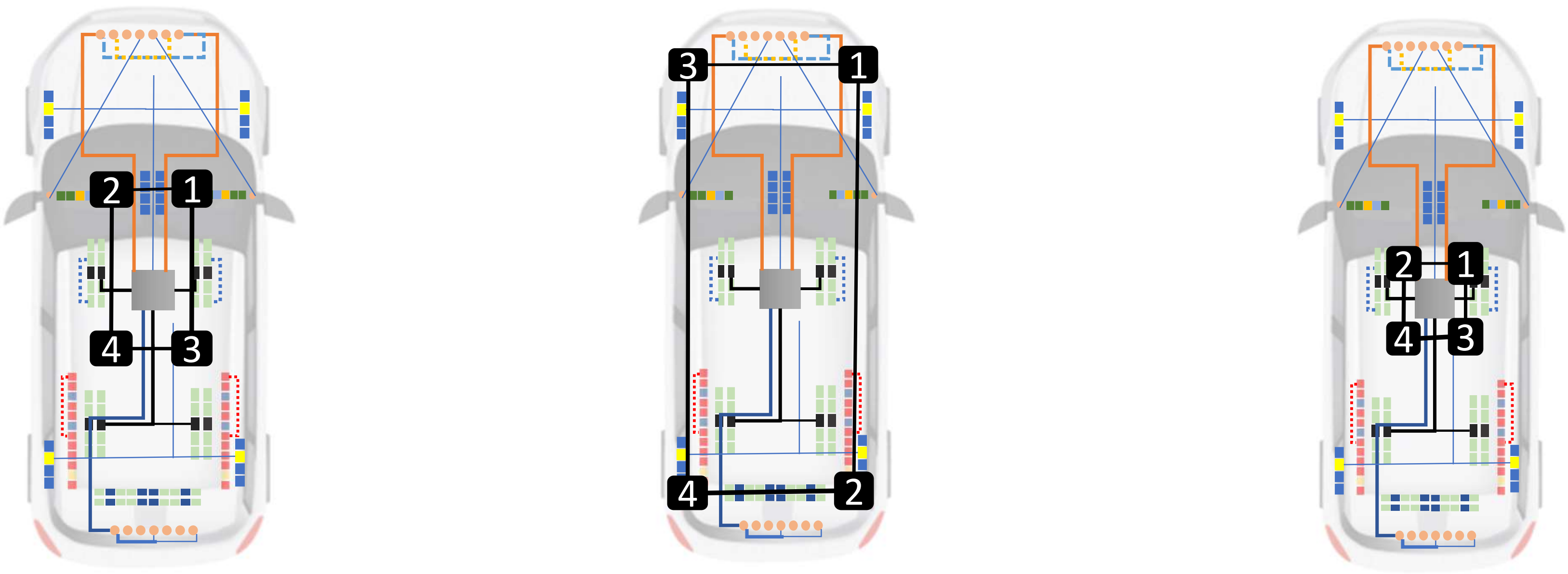
- IRS as a **distributed** service/middleware/library
- A critical application (process) is **fully replicated**
- Replicas form a ***Distributed State Machine*** over a number of ECUs
- Decisions are only made through **Byzantine agreement** (BA/BFT)
- Integrity of decisions is guaranteed despite intrusion faults of ***f* out of *N*** (3*f*+1/2*f*+1) replicas



Need 4 x ECUs?

Leverage modern architectures to host replicas on “similar” ECUs

Component-based vs. Node-based FIT



Domain Distributed

Zone Distributed

Center (clustered?)

The Path to Fault- and Intrusion-Resilient Manycore Systems on a Chip

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

The «YES WE CAN» paper

- ***distributed, parallelized, reconfigurable, heterogeneous…***
  - the very features that cause many of the imminent and emerging security and resilience challenges, can, through ...
- ***replication, hybridization, diversity, rejuvenation, adaptation,***
  - also open avenues for their cure through SoC architecting …

- This disruptive paper (@DSN2023 Disrupt track) suggests paths across the entire SoC hardware/software stack.
- **Modular FIT in modern cars offers a promising application domain**

Shoker, P. Esteves-Verissimo and M. Völp, "The Path to Fault- and Intrusion-Resilient Manycore Systems on a Chip," 53rd IEEE/IFIP DSN Int'l Conference, Disrupt Track (DSN-S), Porto, Portugal, 2023. doi: 10.1109/DSN-S58398.2023.00043.

Distributed Systems-on-a-Chip (DisSoC)

leveraging Ultra-resilient minimal roots-of-trust

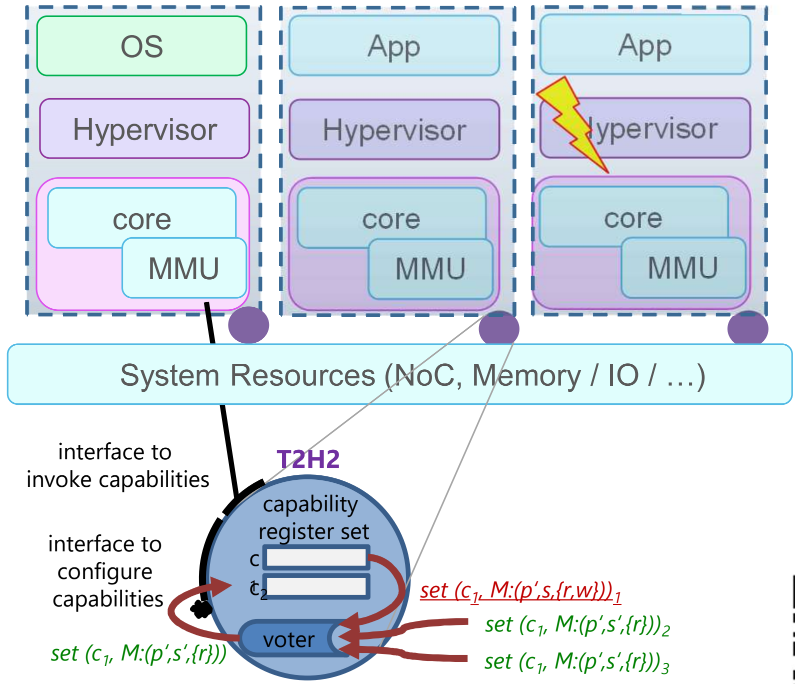
جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of

The Enabling mechanism

=> Patent applications

- Threats have been permeating all levels of architecture.
- And we are always one step "late":
  - we rely on high-level protection (Paxos, BFT,...)
  - threats haunt below (hyp, ME, hw)
  - lost battle: general 0-defect infeasible
- Leverage properties of manycore systems:
  - **distributed systems-on-a-chip (DisSoC)**
  - reinstantiate protection techniques at low enough level (detection, self-check, tolerance)

MIDIR



Behind the Last Line of Defense -- Surviving SoC Faults and Intrusions. Pinto Gouveia, Ines; Voelp, Marcus; Esteves-Verissimo, Paulo. arXiv preprint arXiv:2005.04096 (2020).
Computers & Security, Vol.123, 2022, https://doi.org/10.1016/j.cose.2022.102920.

Savvy: Trustworthy AI/ML powered Autonomous Vehicles Architecture

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

Revisit the current fundamentals of GPT based safety-critical AV architectures, in face of the several problems found:

(i) finding a balance between **intelligence** **and** **trustworthiness**, considering *efficiency and functionality* brought in by AI/ML, while prioritizing indispensable *safety and security*;

(ii) developing an advanced architecture reconciling the **stochastic** nature of AI/ML with the **determinism** of driving control theory

Ali Shoker, Rehana Yasmin & Paulo Esteves-Verissimo. RC3@KAUST. (Work in progress) . Symposium on Vehicle Security and Privacy (VehicleSec 2024) @NDSS Feb. 2024, San Diego, CA-US. arXiv https://doi.org/10.48550/arXiv.2402.14580

Autonomous Driving under attack

"Adversary":

Inadequate or insufficient Machine Learning mechanisms!

Camel visible

No slow down

Tesla hits camel

Ever seen Tesla hit a Camel??

Predicates abstracting the main AI/ML-based AV failure syndromes

- **Issue 1**
  Confusion in Command and Control
  – *(ML model mapping of the controlled process and environment)*

- **Issue 2**
  Better-precise-than-timely (All-or-Nothing)
  – *(ML classification paradigm)*

Incident Analysis (NTSB & NHTSA)
Tesla, Volvo, GM Cruise, Honda Acura

Issue 1
Confusion in Command and Control

Vehicle has not made any slow-down or braking

- AD system could not make a decision
- Late driver handover is being done

Features disabled, ignored sensor inputs

- No reliable system that oversees vehicle state
- No reliable system to take over vs. waiting handover forever

Incident Analysis (NTSB & NHTSA)
Tesla, Volvo, GM Cruise, Honda Acura

Issue 1
Confusion in Command and Control

Vehicle has not made any slow-down or braking	• AD system could not make a decision • Driver handover is being done
Features disabled, broken or ignored sensors	• No reliable system that oversees vehicle state • No reliable system to take over vs waiting handover late or forever
No mentioning to ”invalid” or ”indeterminate” or “not-converging” classification	• ML has not delivered early enough • ML failed to recognize an obstacle

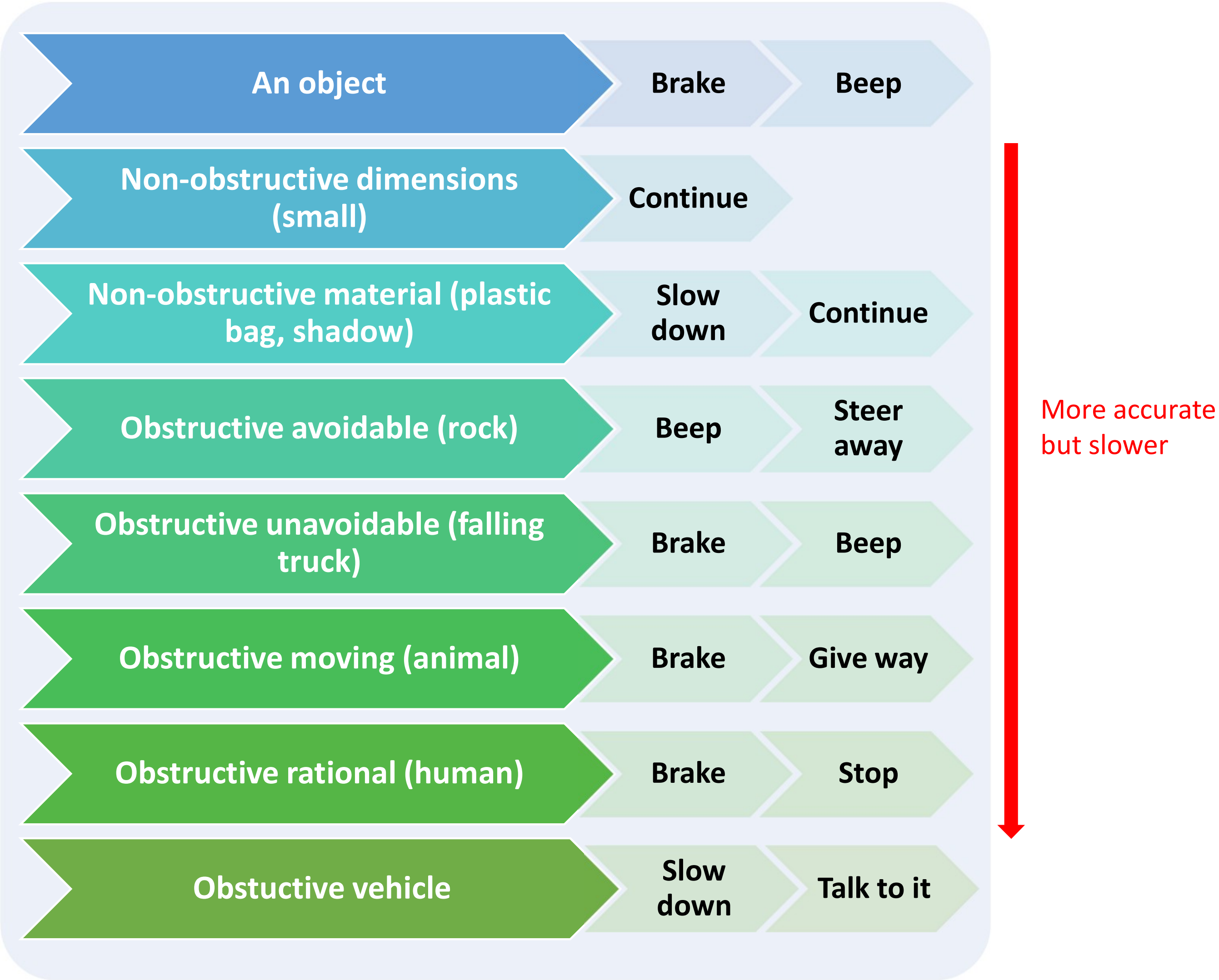
Issue 2
ML classification oriented to Better-precise-than-timely (All-or-Nothing)

Solution Hypothesis

Tune ML to infer useful insights that are **time-bounded**

Dynamic Neural Networks that allow for model deformation using depth and width adjustment (early exiting, skipping, pruning, etc.),
choosing the adequate protocol using Neural Architecture Search or parameter (Weights, Space, or Channel).

Obstacle Avoidance Task



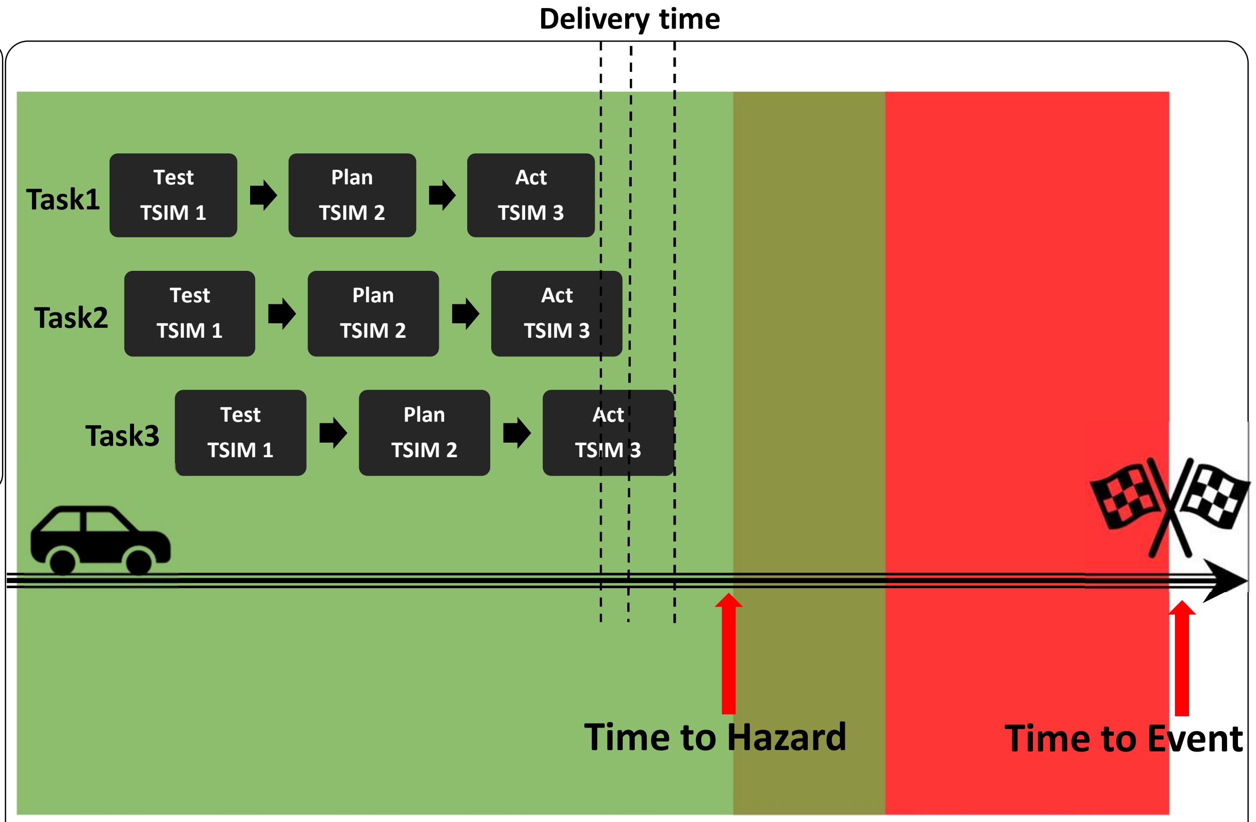
Savvy's approach

Issue 1
Confusion in Command and Control

Solution
Safety-critical Superv. Control System
Hybrid takes-over whatsoever

Issue 2
ML optimized for Better-precise-than-timely (All-or-Nothing)

Solution
ML calibrated for -Time-aware predictive quality degradation



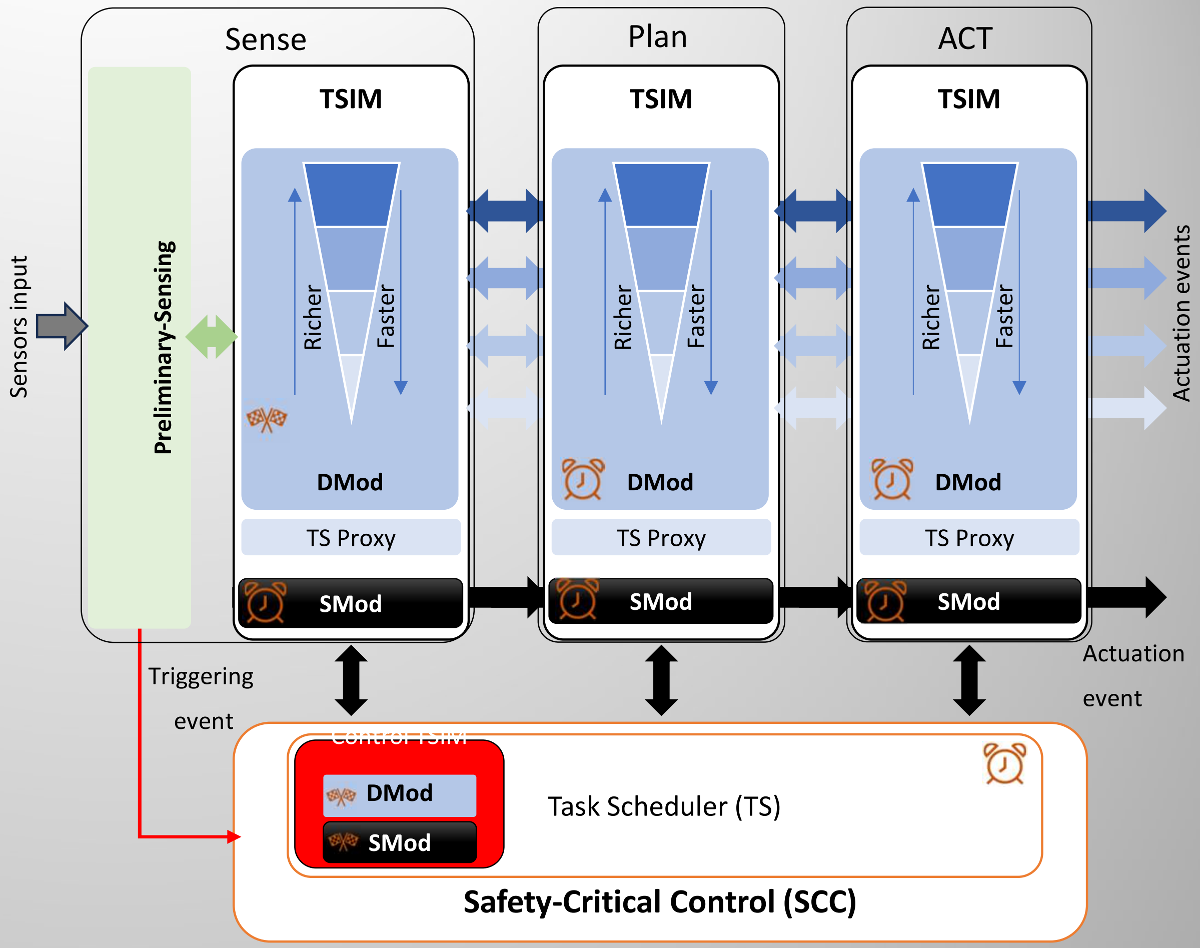
ML deliver here even with ***degraded quality***

Static ***Takeover***

Failsafe ***whatsoever***

Savvy Architecture

- Preliminary Sensing
  - Detect an Event
  - Define Time-to-Event (T2E)
- Safety-Critical Control (SCC)
  - Define Time-to-Hazard (T2H)
  - Set T2E and T2H timers
  - Schedule Tasks over Time-Sensitive Intelligent Modules (TSIM)
- Timer T2H << T2E:
  - TSIM tunes ML model to deliver before T2H
- Timer T2H = T2E
  - Fail-operational: SCC takes over



Crucial non-technical enablers:

- Resilience technologies
  (sustainability through threats)
- Laws and regulations
  (Europe is advanced here)

The NIS2 Directive: A high common level of cybersecurity in the EU

Think Tank
European Parliament

Search...

Home | Research | Events | Infographics | Sources | Partners | Audio podcasts

Research / Advanced search / The NIS2 Directive: A high common level of cybersecurity in the EU

The NIS2 Directive: A high common level of cybersecurity in the EU

Briefing – 08-02-2023

The Network and Information Security (NIS) Directive is the first piece of EU-wide legislation on cybersecurity, and its specific aim was to achieve a high common level of cybersecurity across the Member States. While it increased the Member States' cybersecurity capabilities, its implementation proved difficult, resulting in fragmentation at different levels across the internal market. To respond to the growing threats posed with digitalisation and the surge in cyber-attacks, the Commission has submitted a proposal to replace the NIS Directive and thereby strengthen the security requirements, address the security of supply chains, streamline reporting obligations, and introduce more stringent supervisory measures and stricter enforcement requirements, including harmonised sanctions across the EU. The proposed expansion of the scope covered by NIS2, by effectively obliging more entities and sectors to

C-level executives will be called to order…

Man Who Mass-Extorted Psychotherapy Patients Gets Six Y

April 30, 2024

29 Comments

A 26-year-old Finnish man was sentenced to more than six years in prison today after being convicted of hacking into an online psychotherapy clinic, leaking tens of thousands of patient therapy records, and attempting to extort the clinic and patients.

On October 21, 2020, the **Vastaamo Psychotherapy Center** in Finland became the target of blackmail when a tormentor identified as "ransom_man" demanded payment of 40 bitcoins (~450,000 euros at the time) in return for a promise not to publish highly sensitive therapy session notes Vastaamo had exposed online.

- *Former CEO of Vastaamo, was fired and also prosecuted following the breach. Convicted to 6 months jail, suspended.*
- *The company used username and password "root/root" to protect sensitive patient records.*

Regulation of artificial intelligence
EU AI Act

BBC

Home News Sport Business Innovation Culture Travel Earth Video Live

AI: EU agrees landmark deal on regulation of artificial intelligence

9 December 2023

Share

Topics
European Parliament

Search

Circular economy | Climate and environment | Economy and budget | Gender equality | Workers' rights | All topics

s > Digital > Artificial intelligence > EU AI Act: first regulation on artificial intelligence

EU AI Act: first regulation on artificial intelligence

The use of artificial intelligence in the EU will be regulated by the AI Act, the world's first comprehensive AI law. Find out how it will protect you.

Published: 08-06-2023
Last updated: 19-12-2023 - 11:45
6 min read

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology | Resilient Computing and Cybersecurity Center

TAKE-AWAYS:

Resilient Computing and Cybersecurity Center

Ecosystem mindset

Laws and regulations, "no Far-West"

AV systems (AI/ML or other) cannot ignore distributed real-time systems and control theory

Accidents and attacks, safety and security

Reconciliation of uncertainty with predictability must be an inherent design predicate, not an after thought, a question of "training better"

Modular and technology neutral resilience solutions, from mechanical to cyber world