Workshop program

13.00 – 14.10: **Presentations**

- Workshop introduction, Martin Törngren (KTH) and Viktor Kaznov (Scania)
- **Safety Considerations when preparing for autonomy in the automotive domain**, Masoumeh Parseh, KTH
- Challenges for ensuring functional safety for connected autonomous vehicles, Fredrik Warg, SP
- **Open issues for monitoring architectures**, Jeremie Guichet, LAAS
- **Architecting autonomous vehicles**, Naveen Mohan, KTH
- **Safety Assurance Argument Strategies for Vehicle Autonomy**, John Birch, HORIBA MIRA

14.10-14.30: Break

14.30-16.15: World café sessions with 4 themes:

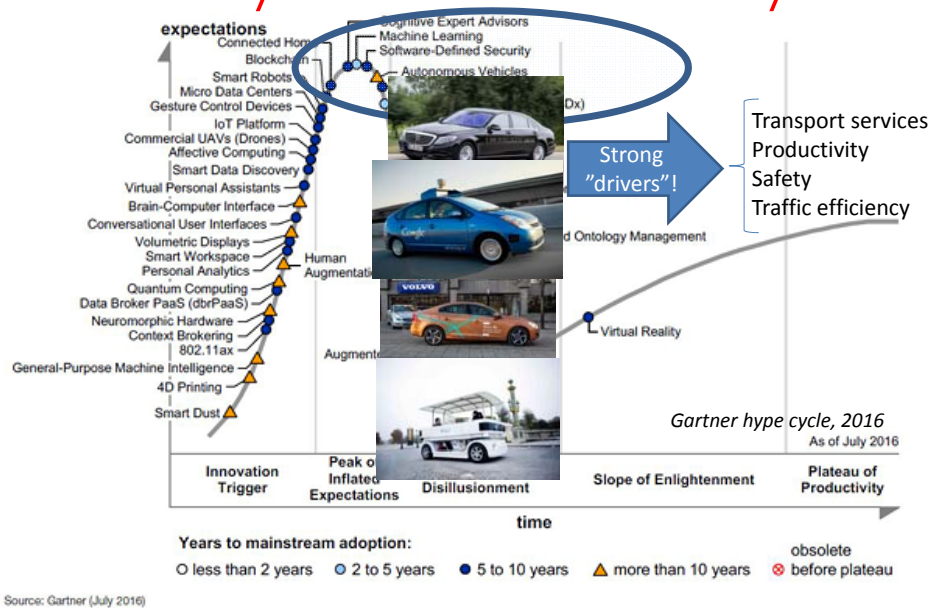
- **Safety analysis** (chair: Sofia Cassel)
- **Supervisor architectures** (chairs: Jeremie Guichet/Lola Masson)
- **Architecting autonomous vehicles** (chair: Naveen Mohan)
- **Safety assurance** (chair: John Birch)

16.15-17.00: Short summaries and wrap up

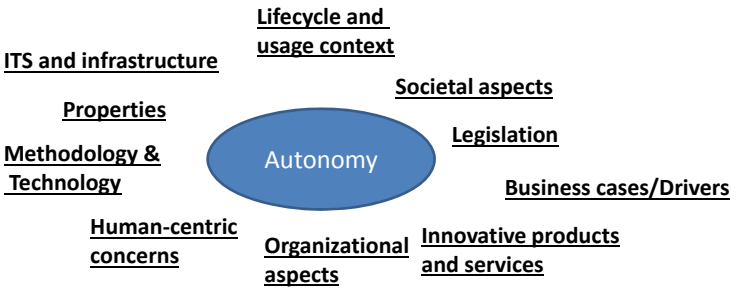
- Summaries per table (by Table chair)
- Wrap-up

Drivers of change and trust!

Autonomy – from fiction to reality

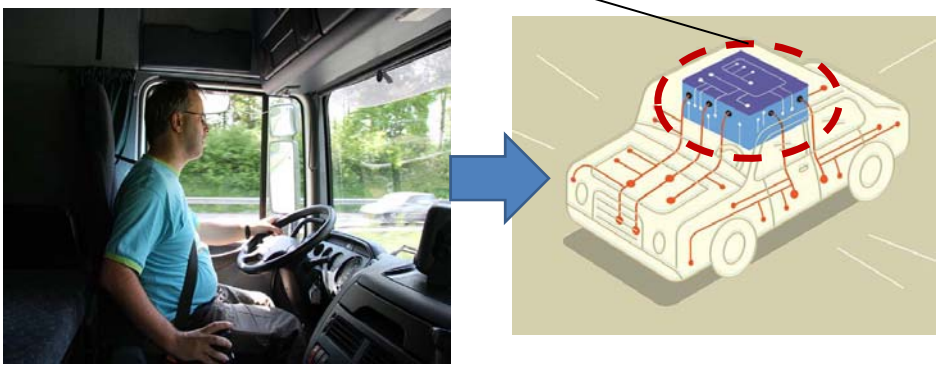


Autonomy related perspectives
- a rich socio-technical area!



Safety related challenges and solutions for autonomous driving?

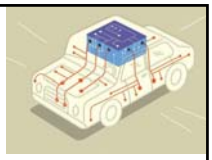
ADI – Autonomous Driving Intelligence



By Veronica538 (Own work)
[CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0) or GFDL (http://www.gnu.org/copyleft/fdl.html)], via Wikimedia Commons

Illustration: Harry Campbell, IEEE Spectrum
http://spectrum.ieee.org/cars-that-think/transportation/self-driving/nxps-bluebox-bids-to-be-the-brains-of-your-car

Outline



- What is special with autonomy in the automotive sector?
- Safety for autonomous driving
  - Key open questions
  - Analysis and promising directions
- Conclusions

"Purely" mechanical vehicle

	Susp	Brake	Steer	Wheel	Diff	Trans	Clutch	Eng	Driver
Susp				X					X
Brake				X					X
Steer				X					X
Wheel	X	X	X		X				
Diff				X		X			
Trans					X		X		
Clutch						X		X	X
Eng							X		
Driver		X	X				X		

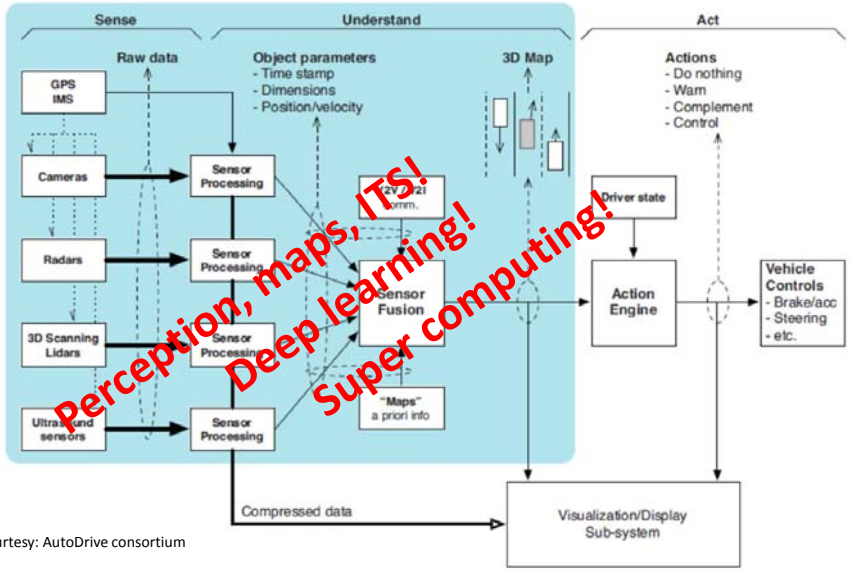
X - Mechanical relations

Fully programmable vehicle!

	Susp	Brake	Steer	Wheel	Diff	Trans	Clutch	Eng	Driver
Susp		P	P	X+P	P	P	P	P	X+P
Brake	P		P	*X*+P	P	P	P	P	*X*+P
Steer	P	P		*X*+P	P	P	P	P	*X*+P
Wheel	X	X	*X*+P		*X*				
Diff	P	P	P	*X*+P		*X*+P	P	P	
Trans	P	P	P	P	*X*+P		*X*+P	P	P
Clutch		P	P		P	*X*+P		*X*+P	P
Eng	P	P	P	P	P	P	*X*+P		P
Driver	P	*X*+P	*X*+P		P	P	*X*+P	P	

P – Programmable relations
X – Possible change

Autonomy related functionalities



Courtesy: AutoDrive consortium

Automotive melting pot

- Electrification and new power sources
- Connectivity and consumer electronics
- Servitization and DevOps
- Automation/autonomy



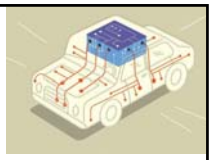

Robotics and AI
System safety, security and dependability
Embedded and high performance computing
Systems and software engineering
Multidisciplinary!

Autonomy in the automotive sector compared to more mature domains?

- Cars in everyone's hand
  - Most complex consumer electronics product
- Largely "uncontrolled" setting
  - "Untrained" users
  - "Unregulated" domain
  - Larger set of usages, scenarios and business cases

  In contrast to e.g. MedTech, Aerospace, …
- Highly integrated systems
  - Bottom-up growth, weak systems engineering

Outline



- What is special with autonomy in the automotive sector?
- **Safety for autonomous driving**
  - Key open questions
  - Analysis and Promising directions
- Conclusions

Key open questions

- Reasoning about and regulating autonomy
- Requirements?
  - What are acceptable levels of risk?
- Safety concerns for advanced perception, planning and control
  - Dealing with AI, uncertainty and complexity
  - Safety practices/standards and architectural concerns
- Autonomous vehicles in Intelligent transportation systems
  - Safety and dependability in Systems of Systems!

Automation levels

Fallback performance by Autonomous driving system!

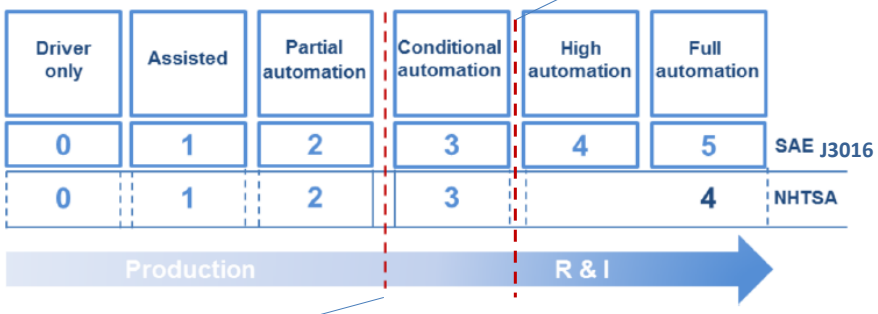


Image source: European roadmap: Smart systems for automated driving

Autonomous driving system monitors the environment

Simplified level concept
Large gaps between levels
Multitude of cases/environments

NHTSA guidelines

NHTSA requests manufacturers to provide reports responding to the guidance (mandatory in e.g. California). The Safety Assessment initially covers the following areas:

- Data Recording and Sharing
- Privacy
- System Safety
- Vehicle Cybersecurity
- Human Machine Interface
- Crashworthiness
- Consumer Education and Training
- Registration and Certification
- Post-Crash Behavior
- Federal, State and Local Laws
- Ethical Considerations
- Operational Design Domain
- Object and Event Detection and Response
- Fall Back (Minimal Risk Condition)
- Validation Methods

https://www.nhtsa.gov/technology-innovation/automated-vehicles

Requirements – acceptable risk level? How much better should "robo-cars" be?

- Fatalities saved vs. People killed by robo-cars?
- Human performance?
  - Approx. 1 fatality per 160 million km's (US statistics)
  - Typical driver crashes once every 257000 km (~ every 12 years, US statistics)
- Societal, ethical, legal, insurance considerations across cases & countries
  - Adjustable risk?!
- Different car maker approaches!?

Safety concerns and implications (I)

- How to ensure that the ADI is "better" than a human driver?
- Testing only will not be feasible!
  - Billions of miles of testing are needed to demonstrate their reliability in terms of fatalities and injuries (Rand Corp)
  - What represents meaningful miles?
- Unlimited amounts of scenarios
  - What are suitable safety analysis techniques?
  - Effective and efficient V&V techniques?

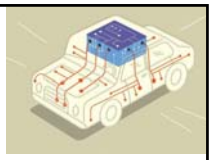
Safety concerns and implications (II)

- Performance and failure modes of machine learning systems
  - Limited transparency and understanding!
- Amodei et al. - Concrete Problems in AI Safety (2016)
  - Extrapolation from limited training data or using an inadequate model
  - Mis-specification of the objective function
    - E.g. negative side-effects
- Stating goals and constraints appropriately
- Robustness considering limited training data

Safety concerns and implications (III)

- Safety approach to deal with probabilistic functions including machine learning?
  - Safety: simplicity; predictability; verifiability
- Applicability of best practices from aerospace?
  - Fail-safe states and separation between safety and main control channels?
- Current automotive platforms and functions typically designed to be fail-silent
- Life-cycle management, repairs, upgrades, security, …

Outline



- What is special with autonomy in the automotive sector?
- Safety for autonomous driving
  - Key open questions
  - **Directions**
- Conclusions

Directions (I): Research and engineering

- Robust/adaptive/self-aware perception and robust AI
- Safety and dependability
  - Safety analysis techniques
  - Run time risk management
  - Cost-effective architectures integrated with planning and supervisory control
- Virtual verification + Formal methods + Testing + DevOps

Directions (II): Safety and availability

	Safety	Availability
Deliberation	Safety related!	High requirements, use-case dependence
Degraded operation	Highly critical safety function (availability)	Required to reach fail-safe state
Reactive/active safety	Highly critical safety function (commission failures, availability)	FS → FO Higher requirements than today

Architectural concepts and considerations:
- Supervisors; Inherent redundancy; integration with existing platforms
- Safety/availability trade-off
- Constraints/decisions for degrading and shutting down

Directions (III): Safety standards evolution

Quest: to provide guidance on how safety-related sensors and perception can achieve the required integrity
- means for reducing risk, validation, acceptance criteria
- beyond guidance in current functional safety standards

- IEC 61508 and IEC 62998 CD
  - IEC 62998 to address in particular guidance for safety-related sensors used for protection of person
- ISO26262 and SOTIF
  - SOTIF to provide guidance regarding special "functional" failure modes, e.g. complex perception

Take aways - the safety vs. smartness challenge for autonomous systems

- Autonomy provides a disruptive change for vehicles
  - Great opportunities for new safe and green systems
  - Strong economical drivers and socio-technical impact
- Challenges
  - Unclear requirements but even more unclear how to ensure that the ADI performs better than humans
  - Dealing with AI, uncertainty and complexity
  - Safety practices/standards and architectural concerns
- Collaboration across multiple stakeholders needed!
- Excellent example of Cyber-Physical Systems evolution
  - Similar challenges will appear in other domains!