

World Café on Autonomy, AI and Self-Awareness

- Key issues in Autonomous Systems, AI and Self-Awareness were elicited in discussion from 3 groups each considering each topic in turn
- The most appropriate means of addressing the issues identified were then discussed in three groups. The issues that could be addressed at Global, European, National and Regional level were then identified – highlighted in red.
- Grand Challenges were identified in the area of autonomous systems and AI

Autonomy, AI and Self-Awareness

- Autonomous Systems
- AI
- Self-Awareness

Key Questions for Future Programmes

- What topics need to be addressed at a European level with coordinated funding?
- What topics would benefit from being addressed at a National/Regional Level?
- Where do topics fit, e.g. CPS, IoT, Big Data, 5G or other labels?

Autonomy (Automation – partial and full)

- Already classification for vehicles – does not correspond to Industry 4.0 – Could be defined at regional or global level
- Trust – Avoid “Hype and hangover”. Can I trust my car?
- How to instil trust? National and European – Different driving styles (Autonomous and Manual – infrastructure)
- Assurance – cannot cover all situations - National (different national traits)- European (Media)
- Mixed-criticality systems – real time embedded hypervisor and separate Cognitive Control with aim of commoditising this for fast track certification
- Dependable communications in networks – bounding of corner cases via probabilistic models
- Current work is based on a common set of models (sandbox) for comparison – question is how do these map to reality – European level
- Interoperability between system components – different approaches from different companies – defining safety boundaries - Global
- Always the case that something will be compromised – build in secure components
- Security – system needs to be resilient – European
- How to get benefits of hype without opening up to new attacks?
- Composability – what should not happen, careful design of components
- Privacy – how to maintain while being resilient – kicking out entities which are misbehaving – National
- Conditional privacy – who has power to remove? National, legislation
- Trade-off between non-functional properties – argue for “smart case” vs safety case
- Emergency braking already exists – system has authority now
- Impact of mistake – what if human makes mistake in developing autonomous function?
- Establish liability - European
- Robotics for Health – National/Regional

AI

- Ethical issues – is there bias? Is it acting in my best interests? Need ethics training for engineers, e.g. Sacrifice you and your car to save pedestrians – European/Global/Needs Regulation
- Transparency – don't know what is going on – European/Global Regulation
- Symbolic AI – go back to 70's – Don't need Google-like AI
- Scientists looking at AI for “general systems”
- New hardware being developed for AI (Deep Learning, Neural Networks)
- Decisions in real time and parallelisation of paths
- More processing, more accuracy but older results – trade-offs in real-time world, cannot sacrifice safety it is a given
- Machine learning to get algorithm to control machine – how to guarantee safety – cannot understand potential failure modes – detecting objects – Recertified/Revalidated European
- Likely to get unwanted behaviour – how to identify and mitigate
- How to guarantee completeness of training set (e.g. handicapped people) – European/Global
- Predictability of behaviour of people interacting with systems
- Car evolving – how to deal with certification if constantly evolving, how are changes OK'd. Idle control already exists – European/Global
- How to limit impact of AI on safety properties
- Probabilistic programming

Self-Awareness

- Situational awareness
- Needs built-in learning – what is it learning?
- Cars – know speed, start sharing data, get shared awareness of each individual, will be deviations – question is how to use information – could be safety risk
- Ice warning – cars ahead
- Data redundancy from multiple sensors
- Temperature sensors in house – can extract other information
- Moving from local to distributed self-awareness
- **Connectivity aspects introduce security issues – National/European**

Potential Grand Challenges

- Definition of classification for general levels of autonomy
- Definition of Ethical basis for AI considering key rules that need to be adopted
- Guaranteeing safety and certification approaches for machine learning algorithms when it is not possible to understand all potential failure modes