Data Department

4 March 2020

The 1st preparatory meeting of HLM CAD Helsinki on 6th February 2020, Schiphol, Amsterdam; Meeting notes

Promoting the transparency of AI systems and algorithms for transport automation

AI and automated driving; (slides attached)

- Eetu Pilli-Sihvola (Finnish Transport and Communications Agency) presented the topic
- Discussion covered for instance trust in AI and cultural differences, different levels of automation and where the focus should be
- The ways of expanding ODDs were discussed as well as the transition from automated mode to driver and the other way around

Setting the scene and the HLM CAD process (slides attached)

Kirsi Miettinen (Ministry of Transport and Communications, Finland) presented the HLM CAD process and the starting points for the work:

- The goal of the HLM CAD process and the Finnish chairmanship is to look into a limited number of possible ways forward in the CAD development. The more concrete the steps are, the better. At the same time, we wish to promote the builduing of the "big picture".
- HLM CAD process aims to build on existing work; no need to start from scratch. The idea is to avoid overlapping work with e.g. the Commission CCAM work.
- All ideas are welcome
- Three focal topics:
    - human-centricity (as regards the issue of algorithmic transparency)
    - enhanced data sharing
    - new regulatory landscape
- Next meetings; listed in the end of the document
- The ministerial high level meeting 6 – 7 October 2020; in Helsinki Save the date notes have been sent out via the Finnish permenant representation in Brussels. Official invitations will follow in March.

Ludger Rogger, DG RTD, European Commission, presented the work of the Commission Expert Group on ethical issues raised by connected and automated driving is currently working on the topic (slides attached).

Geert Van Der Linden, DG Move, European Commission, presented the work of the CCAM Platform.

Transparency/accountability of AI systems and algorithms (slides attached)

- Meeri Haataja (CEO, Saidot) presented the topic. The morning session focuses on the big picture while the afternoon digs deeper into the details.

Questions and discussion:
- Balanced approach between AI accuracy/performance and accountability; there is a certain trade-off between AI performance and its explainability. We need to make choices, and explainability has a high weight.
- AI market approval and lifetime approach; it is not enough that AI systems are approved once as they learn and develop all the time. Targets and allowed changes (boundaries) need to be defined, and if it the change reaches the boundaries, more inspection is needed.
- When it comes to transport; when cars learn while in the market, do they start to differ, or do they learn collectively?

Interactive workshop: Standards and recommended practices for transparency (slides attached)

Meeri Haataja presented the topic and some questions for discussion on metadata and standards:
- Could the IEEE's P7000 process model work in the context of transport automation
- Could this kind of a metadata model / standard provide common ground and efficiency for the industry in transport automation
- How could standard model facilitate regulation and smooth oversight going forward?
- What are the industry specifics, which should be taken into considetation, while building an industry specific version of transparency metadata model?

Discussion:

Transparency and explainibility
- We need transparency for acceptability - no automated cars can be allowed on the road if we do not know what they are. For the moment there are no tools for transparency in any sector, neither in transport automation.
- Details are important but who needs to know and to understand them? Consumers might not want to know but someone else, e.g. ethics regulatory supervisor.
- Work is going on in UNECE WP.29 and at the European Commission on automated driving and the approach that the automated vehicle has to obey traffic rules and how to achieve this performance. WP.29 tried to define transport safety and accepted level: vehicle shall not cause any non-tolerable risk, so there is some leeway.
- Can we define how much more safe automated driving is compared to human drivers? What if we cannot explain it anyway?
- There is a need to assess on a statistical level. As accidents happen rarely, there is not so much data.
- A self learning system is not as easy to explain as a gear box
- We need algorithms to explain algorithms. It is hard to do at a general level but has to be done industry-specific and we need concrete tools for that. We need sandboxes where industry and regulators can explore this together, in the right context.
- Explaining problems in driving has proven useful – not only the consequences we wish to avoid. E.g. mitigating risk by lowering speed instead of just focusing on avoiding collission.
- Local and global explainability were discussed. Should the designer of the algorithm be able to explain every detail? Does it depend on the context? An analysis is expected on any fatal road accident – can we have absolute explainability? Explainability means that we try to answer as

Ministry of Transport and Communications

Office
Eteläesplanadi 16
FI-00100 Helsinki

Postal address
PO Box 31
FI-00023 Government
Finland

Telephone
+358 295 16001

www.lvm.fi
firstname.lastname@lvm.fi
kirjaamo@lvm.fi

many questions as possible but there are different levels. We cannot explain everything with humans either.
- Difference between explaining to users vs. to experts.
- There is always some risk but the system has to manage the risk.
- Risks can be defined in many ways, we need to sum costs and risks. With the help of AI the total risk needs to be smaller.

Simulation
- How to assess automated driving in a similar way as human driving, do we need a driving test and what kind? How much more do we require from automated vehicles compared to human drivers?
- The key difference is the scale; individual driver vs. automated system.
- Do we need to accept that automated driving also causes casualties, like road transport now causes, and yet we do not stop driving?
- Simulation is needed and we need to think about the possibilities of simulation, in a safe way. Some problems; for instance testing on public roads would require knowing before hand how it will go on public roads.

Cybersecurity
- Important to take into account cybersecurity.
- Safety and security are the key drivers why we need transparency, safe and secure by design approach. "By design" approach. To prove those features, we need transparency. Explainibility is one element, so everyone can understand how decisions are made.

Interactive workshop: Certification (slides attached)

After a presentation by Meeri Haataja, the participants discussed the theme in two groups, focusing on the following questions:
- Could transparency certifications be used as a means for regulation in the space of transport automation?
- Could the presented IEEE's ECPAIS certification scheme provide a good ground for industry tailored version?
- What are the industry specifics, which should be taken into considetation, while building an industry specific version of transparency certification?
- What kinds of next steps could we take in order to take industry tailored certifications forward?

Discussion:

- Transparency certification could and should be part of the regulation. What kind of certification? Technically it could be built in the regulation system. Is it practical?
- Transport is a high-risk sector, so it needs certification. Industry specific: Does AI need driving licences? High-level meeting could address the gap between technological development now vs. certificates etc. needed for AI driving licence?
- Certification is also a part of the Commission CCAM platform research agenda. CCAM has several different research areas and it provides a forum for working together.
- Also social implications of certain AI or algorithmic systems were touched upon
- Data sharing, test data sharing
  - Data sharing is the key
  - How can we enhance data sharing? While willing to contribute to mutual benefit and common good, there might be lack of money and trust
  - Public sector in many countries is already obliged to make data available and public

- o Data sharing is important for testing purposes. Individual datasets for testing can be biased, so test data needs to be shared and thus the testing provides more accurate results.
  - o Virtual testing; worries over biases and training/testing data.
- Testing frameworks
  - o All testing frameworks have different criteria, including reporting criteria.
  - o Should they be somewhat harmonised? Is this something that the high level meeting could address on the political level? The topic was discussed also at the high level meeting in 2017. Lot of work has been done but perhaps the time was not right yet. Back then it was pointed out that it its hard to set a minimum level by law. However, different frameworks could be compared and perhaps a mutual recognition system could be established. Same issue has been also discussed at CCAM platform WG1.
  - o After all there is a limited number of testing frameworks so harmonisation to some extent seems possible
  - o Many countries could be greatful for the possibility to compare and to learn from best practices
  - o Testing happens in different environments, howe to compare highways and cities or real and virtual kilometres?
- IEEE certification activities
  - o The work is led by the industry and based on voluntary work. The IEEE organisation provides tools for the work. There needs to be enough participants.
  - o As regards transport automation and AI, it could be possible to establish a new group. During the very active phase of the work, it could take several teleconferences a week etc.
  - o For the moment the work is not sector specific - transport sector has not been addressed yet.
  - o THE HLM could address the cooperation with IEEE as one possible example.
- Global cooperation:
  - o UNECE provides the global cooperation forum.
  - o Global standards are imporant while at the same time there is competition. Data sharing on a global level does not seem feasible.
  - o Perhaps political message is needed to highlight the importantance of global focus - focus on the EU context is not enough.

Closing remarks and next steps

Following meetings:
- 17.3.2020: 2nd preparatory meeting (Schiphol). Theme: Enhanced data sharing
- 27.4.2020: 3rd preparatory meeting (Helsinki). Theme: Regulation of transport automation
- 11.6.2020: 4th preparatory meeting (Schiphol). Theme: 1st draft of the outcome of the HLM CAD
- 10.9.2020: 5th preparatory meeting (Schiphol). Theme: 2nd draft of the outcome of the HLM CAD
- 6.-7.10.2020: Ministerial HLM CAD in Helsinki