Position Paper on:

Big data, Artificial Intelligence and their applications

In combination with Big traffic Data, Artificial Intelligence (AI) techniques play a critical role in the development of automated driving technologies. Big Data feeds the machine learning algorithms that drive the continuous improvement of AI-based CAD functions towards higher-level automation. On the other hand, AI techniques are essential for analysing and annotating the collected Big traffic Data and converting it to useful information that can to be used in the development and validation of CAD functions.

Impact

The world of Big Data and AI applications is not only rapidly growing but also very dynamic. The push towards increasing levels of automation of automated driving systems forces the market to accelerate the development of new automated driving services, activities and applications including (but not limited to):

- Analysis of CAD systems: driving patterns can be identified through the development of Big Data analytics tools (for extracting meaningful information from mixed data), which will lead to improved insights, e.g. in driving and recharging patterns, driver-behaviour, understanding the impact of CAD for safety, comfort and mobility in general.
- Validation of AI functions in CAD systems through the definition of scenario-based assessment methodologies, including the classification of real-life scenarios, development of CAD functions, safety monitoring. Furthermore, AI techniques accelerate the software development and product cycles.
- Improving situational awareness using AI techniques e.g.:
    - Improved on-board sensor fusion to estimate the relative position and velocity of the surrounding vehicles - improved camera and radar fusion (see also CARTRE Theme on In-vehicle enablers)
    - prediction models for other road user behaviour (especially complex behaviour like pedestrians)
    - building accurate maps for on-road concurrent mapping and navigation
- Improve the quality of the performance of CAD functions by self-improving mechanisms, possibly shared across cars.
- Open new fields of applications such as taxi services, car sharing or find-a-parking-spot services.

Challenges

These development processes need to find solutions for several challenges in different domains and research areas. It is expected that the challenges that are encountered until 2020 will most likely still carry key uncertainties in 2040. These challenges can be arranged under four different categories:

Technical challenges

- What CAD data is most valuable to share?
- How to handle the reducing validity of Big traffic Data over time?
- How to select a training data set for validating specific AI CAD function?
- How to assess the completeness of the scenarios used in training and validating an AI function

Policy challenges

- How to overcome privacy and security barriers for sharing big data?
- What is an acceptable distinction between company sensitive data, personally sensitive data and research relevant data?

Organisation ecosystem challenges

- How to handle different data sharing needs across various stakeholder types?
- How can we share the investments on developing AI functions?

User acceptance

- How much safer must an AI driver be compared to human drivers to foster user acceptance?
- How should the difference in media coverage for human and AI errors be approached?

Input for the EU research agenda

There are various challenges in sharing and using Big Data and AI, both in the short and long term. The CARTRE working group concluded the following focus points for the EU research agenda:

- **Data sharing:** Policy and regulations for data ownership are strongly needed to stimulate and encourage car manufacturers, other institutions and even individuals to share driving data.
  Very diverse opinions exist on who owns the data. It is unclear whether data ownership may be regulated or driving data may become open within privacy, security and confidentiality constraints. This has a great impact on the re-usability of data that has been collected or aggregated. The importance of data sharing is underlined by its relevance for data-hungry machine learning algorithms.
- **Privacy and security:** Policy and ethics on the appropriate use of driving data are needed.
  Big data collection has a significant impact on privacy and security. Current policies are fragmented and were not prepared with big data in mind. The willingness of users to accept limited privacy may also increase, if a clear returned benefit is perceived.
- **Regulations, ethics and insurance:** A framework adapted to the use of AI functionality is required.
  The shift from human to AI vehicle control poses fundamental questions concerning regulations, ethics and liability, e.g. concerning the programming of AI behaviour for accidents and the subsequent determination of the guilty party
- **Data storage and accessibility:** A framework for storing and accessing shared driving data must be defined.
  Data is collected from many different sources, by different parties and in different formats. Technically the storage and accessibility of this data will become a real challenge at some point. This may require new data storage solutions.
- **Assessment and validation of CAD functions:** There is an urgent need for a harmonised framework for the validation and assessment of CAD functions.
  This includes the harmonisation of test cases and training data sets for AI functions and the question of completeness of training scenarios.
- **Investment for developing CAD functions:** Encourage and enable sharing the investments on developing AI functions for automated driving.
  Currently, massive parallel investments are made in machine learning and AI technologies. Sharing investment will accelerate these developments and naturally works/leads towards harmonisation.