



COUNCIL OF EUROPE



CONSEIL DE L'EUROPE

Strasbourg, 19 May 2019

CDPC(2019)8FIN

EUROPEAN COMMITTEE ON CRIME PROBLEMS (CDPC)

QUESTIONNAIRE CONCERNING ARTIFICIAL INTELLIGENCE AND CRIMINAL JUSTICE (using the example of Automated Driving)

Document prepared by the Working Group of Experts on Artificial Intelligence and Criminal Law, chaired by Mrs Sabine Gless, Professor of Criminal Law and Criminal Procedure, Law Faculty of the University of Basel

Introduction

A. Background

In the 21st century technology is rapidly evolving and in recent years has been noticeably driven by the use of Artificial Intelligence (hereafter AI). Long-term technological trends in this domain suggest that various forms of AI will become more and more involved in modern civilian life by operating and engaging in co-operation with humans. The increased presence of AI in everyday life and in various parts of the criminal justice system presents challenging questions to the Council of Europe as a pan-European organisation (see Council of Europe general activities on AI on <https://www.coe.int/en/web/artificial-intelligence/home>) and to all its member States. Domestic legislation has not always addressed the issue in a systematic way. However, more recently, some countries have adopted specific regulations and certain member States have made substantial progress in their national legislation on driving automation while some have even adopted statutes explicitly governing liability for correct use for the intended purpose.

The Council of Europe committee on criminal problems (CDPC) (<https://www.coe.int/en/web/cdpc/home>) started its work on AI and criminal law in 2017 and prepared a first document on this topic: a concept paper on “Artificial intelligence and criminal law responsibility in Council of Europe member States - the case of automated vehicles” (<https://rm.coe.int/cdpc-2018-14rev-artificial-intelligence-and-criminal-law-project-2018-/16808e64ad>). This document (hereafter the Concept Paper) contains the main elements of a project to be implemented over the next few years by the CDPC.

On 28 November 2018, the CDPC organised a Thematic Session on AI and criminal law responsibility where driving automation served as an example for situations where pervasive computing is responding to human needs, the main objectives of which were to:

- i. Examine and ascertain the current existing scope and substance of relevant national criminal legislation and international law, using automated driving as an example for AI deployment, as well as determine where and how regulatory powers are established within the competent national public authorities.*
- ii. Determine where certain conduct has been or should be prohibited and criminalised in relation to the delegation, division or assignment of tasks, functions and behaviours to automated technologies, and the possible cross-border relevance.*
- iii. Illustrate the findings under ii (see supra) using the case of automated driving: should new principles and norms of attribution and accountability for natural or legal persons be established to uphold Council of Europe Conventions’ goals if automated driving (or other Artificial Intelligence deployment) operates across borders.*
- iv. Examine the scope and substance of an international legal instrument to provide common standards for the criminal law aspects of automated technologies, in particular automated vehicles.*

As follow-up to this Thematic Session, the CDPC set up a working group of experts representing member States supported by some scientific experts (hereafter the working group) and tasked them to assist the CDPC in implementing the project activities contained in the Concept Paper. The working group held its 1st meeting on 27 March 2019 and prepared this questionnaire, which is the first output as foreseen in the Concept Paper (see Output 1 on page 8)¹.

¹ **Output 1 (excerpt): 5.1.1 Research project on national criminal law and international legal framework: (a) Activity:** A questionnaire followed by a compilation of responses and analysis; **(b) Reasons:** In order to survey the current regulatory framework for AI, and in particular automated vehicles, key national-level information should be extracted from the member States; **(c) Working methods:** A questionnaire is to be developed and distributed to the relevant ministries (or other entities, as appropriate); the answers will be analysed by an expert or panel of experts.

B. Objectives and scope

Driving automation is a poignant example of **narrow AI**, a bundling of certain techniques already enabling human co-operation with driving systems, i.e. **(ro)bots** that can interact with each other as well as with a human user in order to (temporarily) take over the driver's tasks as a first step. The eventual goal is autonomous driving cars.

The fact that an industry standard distinguishes between different levels of **driving automation** ([Norm SAE J3016_201401](#)) facilitates a comparative approach to the possible impact for criminal law, criminal procedure and mutual legal assistance. Of interest here are the transitions from level 2 onward to level 3 and 4 and possibly to level 5. At level 2 a car can execute dynamic driving tasks but the driver must monitor and overrule the system if necessary. At level 3 a driver no longer needs to monitor when the system is activated, but the driver must respond to a takeover request. Level 5 envisages autonomous driving without a human driver. During automated driving, data is automatically generated that could offer relevant information in a criminal trial after a traffic incident.

While driving automation is probably the most prominent example of AI-human co-operation in daily life, other fields are gaining importance (such as medical devices or service robots). The common feature is the ability and necessity to take in information, react and learn from "experience" without human interference. Therefore neither producers and programmers nor users can foresee *all* possible actions of an AI-driven (ro)bot. This means that one cannot reduce to zero the possibility that such a device may also cause harm to others in a particular situation. Among other things, this fact suggests two mutually-exclusive conclusions as to liability for negligence. It could be argued that no one can be held responsible because the machine is acting "on its own"; alternatively, it could be claimed that a producer can foresee harm and therefore should face *de facto* strict liability for the results of a robot's acts.

The CDPC wishes to assist member States with a common approach to a regulatory framework when the many beneficial, yet potentially risky, uses of AI are to be integrated into daily life. One important aspect for criminal justice systems is accountability for harmful consequences. Duly considering the *ultima ratio* of criminal regulation in this complex field, this project focuses only on situations where the level of harm, or the seriousness of the violation of the obligation breached, could or should entail criminal responsibility, and where the use of AI affects criminal justice systems.

C. Example case

This following scenario illustrates some of the challenges arising for criminal law, criminal procedure and mutual legal assistance from ambient intelligent environments (using the example of driving automation):

Imagine that, for the first time, a vehicle equipped with an "autopilot system" can be used legally on highways in your country. The automated driving system must be used in harmony with the authorisation which requires – among other things – that the human driver is ready to take over the steering wheel within 20 seconds. To ensure the driver's fitness to take over, the producer installs a drowsiness detection system monitoring the driver (seating position, face and especially eye movements) and stores the data with a cloud service provider. During the first months of operation of such cars, it turns out that a certain weather phenomenon in your country (be it morning mist, a sandstorm, midday sun or garbage thrown on the roadside) triggers faulty reactions in the driving assistant's system – especially false-braking, i.e. braking for the wrong reason, for instance a plastic bag drifting in the wind. The producer and all component suppliers do their very best to fix the problems. However, it is clear to everyone involved that the cars will need time to adjust to particular local conditions.

Questions *(When answering the questions you may tick more than one box.)*

1. Does your national legislation and/or case law specifically **address criminal liability issues connected to driving automation**?
 - a) If so, could you please:
 - (1) provide, if available, the relevant texts (in English or in French);
 - (2) indicate whether criminal responsibility is allocated to a specific person (natural or legal, e.g. *driver, producer, programmer, fleet supervisor, tele-operator etc.*) and on what standard it is based (*i.e. strict liability, negligence, intention*).
 - b) If not, will general rules apply in a situation where a driving assistant /"AI" steers a car when the accident happens, and what kind of problems are to be expected?

2. Does the lawmaker in your country plan **legal reforms** with regard to (criminal) liability connected to **driving automation** (Level 3 of driving automation, see supra A.)?

3. Does the lawmaker in your country plan **legal reforms** with regard to **autonomous systems**, e.g. autonomously driving cars on public streets (without a driver present, *Level 5 of driving automation*, see supra A.), like granting an "e-personhood"?

4. Does your law **differentiate criminal liability for lethal incidents** (on public roads) depending on:
 - a) **gravity of the breach of law** (e.g. intentional killing, gross negligence, negligence)

If so, is a distinction made based on:

 - general criminal law (applicable to all offences, e.g. intention, negligence)?
 - specific rules on criminal liability for death caused involving the use of a vehicle, i.e. death by dangerous driving, vehicular homicide)?
 - other? (please explain briefly)
 - b) **perpetrator**, if so is a distinction made based on the following categories:
 - driver of a car?
 - owner, "keeper", "registered user", "registrar"?
 - producer (e.g. corporate manslaughter)?
 - other? (please explain briefly)

5. According to your domestic law and/or case law what does **criminal negligence** resulting in harm require:
 - a) gross violation of a duty of care?
 - b) foreseeability?
 - c) preventability?
 - d) (solely) a violation of a duty of care?
 - e) recklessness?

If necessary, please provide a short answer.

6. Does your law use the concept of **strict liability** :
- a) in criminal law, i.e. “punishment without guilt”?
 - b) for (administrative) traffic offences, e.g. handing parking tickets to the owner of a car?
 - c) any other ? (please specify).

Imagine in the example case (see *supra* C.) unfortunately the car hits a human while driving on autopilot and the person dies. It can be established that the car’s sensors were defective, but also the braking assistant had a severe software defect.

It is however impossible to prove which fault caused the accident. Does your national legislation and/or case law address this problem of a criminal “**contributory negligence**”?

If your system addresses the problem, is it seen as a problem of

- a) theories of causation?
 - b) “complicity” (or rather collaboration in negligence)?
 - c) any other? (please specify).
7. Imagine further that it could be proven that the car’s sensors did not pick up the victim, most likely because he/she held a bag at arm’s length and the engineers had “tuned out” bag images from the sensors’ vision in order to prevent “false braking”. In such a case, criminal justice systems may provide an option to forgo criminal prosecution, arguing that in the light of the overall social benefits a particular type of risk taking should not be punished even if harm is caused as long as the person in question does its best to comply with all requirements of safety and security. (Such a notion, for instance, is prevalent when using airbags in cars where although there is a minimal risk that this safety device might open because of a pothole and kill a passenger, it will moreover save lives in many situations.)

Does your domestic law recognise the notion of a “**societal accepted risk**”?

If so, how? (please explain briefly).

8. Does your domestic legislation and/or case law address problems of the obtaining, presentation and evaluation of **digital evidence generated during driving** (e.g. evidentiary requirements; requirement for the car producer to provide data, in readable format)? If so, please explain briefly.
9. Are there **problems with digital evidence** stored in a car or at the manufacturer due to
- a) encryption?
 - b) data storage with a cloud service provider or abroad in a third country?
 - c) technology design of data generation?
 - d) any other? (please specify).
10. Imagine in the example case (see *supra* C.) that the driver overrules the drowsiness detection system’s suggestion to take a break, because he/she feels awake and a lethal accident happens and the prosecution wants to introduce the drowsiness alert as evidence. In your legal system, could the driver challenge the **credibility of the system or the reliability of such evidence** (possibly even in a way that is similar to challenging the credibility of a witness and the reliability of a testimony)?
11. Do authorities in your jurisdiction use **digital analytical tools** to enforce safety on public roads (i.e. use of digital breathalysers or predictive policing/**profiling** to identify high risk drivers)?