COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

Strasbourg, 16 October 2018 CDPC(2018)14Rev
cdpc/docs 2018/cdpc(2018)14

EUROPEAN COMMITTEE ON CRIME PROBLEMS (CDPC)

CONCEPT PAPER

PROJECT TITLE:	ARTIFICIAL INTELLIGENCE AND CRIMINAL LAW RESPONSIBILITY IN COUNCIL OF EUROPE MEMBER STATES - THE CASE OF AUTOMATED VEHICLES
PROJECT AREA:	Council of Europe member States
BUDGET:	Approximately seven hundred and sixty-five thousand euros (765 000€)
DURATION:	2 years
IMPLEMENTATION:	Directorate of Information Society and Action against Crime – DGI, Council of Europe

Table of Contents

Council of Europe Project on Artificial Intelligence (AI)[1] and Criminal Law Responsibility

1. Problem analysis and need assessment

After some incidents with automated vehicles in Council of Europe member States[2], and beyond[3], the following question has been raised: who will be responsible if a completely automated driving vehicle[4] injures or kills a human? With self-learning algorithms driving a car, the more general question arises: how should criminal law address Artificial Intelligence (AI)?

The long-term trends in technological development, and the case of automated vehicles, suggest that AI and machines with autonomous functionality will become ever more present in advanced societies, and that States thoughtfully need to consider how to deal with this in their legal and regulatory framework. A first essential step has already been taken with the introduction of technical standards for special permits allowing automated driving in domestic jurisdictions, but some jurisdictions take the position that no permission is needed for activity that is not considered to be illegal.[5] As vehicles quite frequently cross borders, it appears to be in the interest of the member States of the Council of Europe to foresee how its standards could be adapted so that to ensure their co-operation in future cases, should automated vehicles cause accidents in other countries or illegal activity affect more than one jurisdiction. Unfortunately, it

[1] There is no agreed upon definition of Artificial Intelligence (AI), but for the purpose of this paper the Council of Europe recognises the term as encompassing systems that are operational and capable of performing complex tasks whose goal is to achieve the imitation by a machine of the cognitive abilities of a human being. https://www.coe.int/en/web/human-rights-rule-of-law/artificial-intelligence

[2] For instance accidents involving assisted driving in Germany (AG München, Urteil vom 19. 7. 2007 - 275 C 15658/07), Norway <https://newatlas.com/tesla-autopilot-fema/46045> or Switzerland (https://www.nzz.ch/panorama/tesla-fahrer-will-nach-unfall-milderes-urteil-ld.1334364) as well as https://www.youtube.com/watch?v=qQkx-4pFjus or autonomous driving in Switzerland https://www.swissinfo.ch/eng/autonomous-post-bus-gets-in-accident/42467476.

[3] The Dutch vehicle approval authority RDW apparently has asked the United States National Highway Traffic Safety Administration (NHTSA) for details after a fatal Tesla crash, to see if cars equipped with the autopilot function, approved in Europe by RDW, are safe http://fortune.com/2016/07/14/tesla-crash-netherlands

[4] According to the SAE standard J3016_201806 <https://www.sae.org/standards/content/j3016_201806/> one differentiates Six Levels of Autonomy, starting with Level 0: No Automation, Level 1: Driver Assistance, Level 2: Partly Automated Driving Level 3: Highly Automated Driving, Level 4: Fully Automated Driving, Level 5: Full Automation (Driverless).

[5] See e.g. the information sheet for exemption permits provided by the Swiss Federal Roads Office FEDRO/ ASTRA <https://www.astra.admin.ch/astra/de/home/themen/intelligente-mobilitaet/pilotversuche.html>.

remains highly unlikely that the risk of accidents will drop to zero.[6] It is also foreseeable that some individuals will maliciously use artificially intelligent devices to carry out criminal offences[7].

AI and automated vehicles are mostly being used in a restricted, controlled circumstances at the time of writing. Among other things this is due to the fact that machine learning can be implemented in different ways, and European countries have opted for a "slow approach".[8] However the increased presence of AI in civil life presents a set of challenging questions for legal systems across Europe. Although there is an inherent unpredictability to these trends, current forecasting suggests that over the next five to ten years automated vehicles, for instance, will become much more prevalent in daily life, transport and industry,[9] and while promising substantial safety benefits, will not prevent all accidents. The benefit of establishing clear, common rules of criminal liability will benefit a proper administration of justice.

The relatively simple question of *who* is to be held criminally liable for harmful consequences as a result of a machine's autonomous decision-making processes does not always have a simple answer. For in criminal law it is difficult to deal with "criminal behaviour" of non-human beings; if AI takes the place in the driver's seat there could be a responsibility gap. One of the fundamental aims of this project, and its potential outcomes, is to assess the need to put in place consistent regulations among States to determine situations of criminal liability, in particular in situations of accidents causing serious damage, and thereby avoid undesirable effects on the safe use of these advanced technologies, and to prevent possible adverse impacts.

While there have long been ethical debates in academic research and speculative fiction regarding the potential benefits and dangers of artificially intelligent machines, there has been comparatively little institutional analysis of how to realistically resolve the specific criminal liability issues that are likely to emerge in the coming years. The question of criminal responsibility illustrates this: the legal framework currently applicable to the development and utilisation of automated vehicles (or other AI deployment) is based on normative principles developed during the pre-digital era. As a result it is unclear in various situations as to how and when responsibility for harm can be determined. In the interests of ensuring adequate means of accountability for situations where automated vehicles (or other AI deployment) may cause harm to a human being, it is necessary to help establish a clear criminal-law frame. It could be therefore valuable to set up rules governing any potential criminal liability in advance to ensure that in cases such

---

[6] Accidents are still happening: Daisuke Wakabayashi, "Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam" *The New York Times* (19 March 2018), available at https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html
Bryant Walker Smith, 'Automated Driving and Product Liability', *Michigan Law Review* (2017), available at: https://digitalcommons.law.msu.edu/cgi/viewcontent.cgi?article=1187&context=lr

[7] https://www.theguardian.com/technology/2016/sep/20/tesla-model-s-chinese-hack-remote-control-brakes; https://www.bleepingcomputer.com/news/security/volkswagen-and-audi-cars-vulnerable-to-remote-hacking.

[8] https://www.bloomberg.com/news/articles/2018-03-20/it-s-a-good-thing-europe-s-autonomous-car-testing-is-slow

[9] For example, IHS Markit forecasts that millions of cars "with some form of autonomy" will be produced and sold over the coming decade, with the market potentially reaching 600,000 vehicles in 2025, and potentially up to 21 million vehicles sold per annum in 2035. Similarly, industrial robotics is also expected to rapidly rise, as the total sales of smart autonomous machinery increase by an average of 12% per year worldwide.

as a car collision or a drone crash, no State will have to face an unclear legal situation due to unsuitable or out-of-date rules. This project also has for aim, considering the *ultima ratio* nature of criminal regulation in this complex field, to address the circumstances in which the degree of harm or the importance of the obligation breached could, or should, engage criminal liability.

As the potential widespread adoption of automated vehicles will affect all Council of Europe member States and beyond, there is a role for the Organisation to play in facilitating the general development of the principles pertaining to AI deployment. As far as the more specific issue of criminal law responsibility is concerned, the European Committee of Criminal Problems (CDPC) of the Council of Europe can help States to elaborate common legal standards providing an adequate, comprehensive and straightforward regulatory system that while recognising the many beneficial uses of automated vehicles will also guarantee a clear framework to address the possible abuse and harmful consequences of AI. In order to maintain good co-operation in criminal matters among the members of the Council of Europe several issues should be addressed including the question of how different approaches in testing and using automated vehicles can translate into "permissible risks" not criminalised in domestic law (like the different uses of technologies in cars) as well as the question of whether an automated vehicle may eventually have to answer the law as an e-person (similar to corporations as legal persons) or whether criminal justice is for "human persons" only.

This process should involve a number of actors, including, for instance regulatory authorities such as transportation ministries or road safety authorities, and others who are developing and implementing safety standards and procedures to determine regulatory compliance for automated vehicles.

The technology standards developed at international level[10] could then pave the way for suitably careful legal regulation at national level relating to the use and oversight of automated driving including principles for allocating criminal law responsibility for the employment of AI and appropriate sanctions and measures, where necessary, as well as a common legal framework to resolve any transborder issues and a legal basis to facilitate mutual legal assistance in criminal matters.

[10] See e.g. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles J3016_201609 < https://www.sae.org/standards/content/j3016_201609/>, referred to in A common EU approach to liability rules and insurance for connected and autonomous vehicles <www.europarl.europa.eu/RegData/etudes/STUD/2018/615635/EPRS_STU(2018)615635 >, p. 42; The International Organization for Standardization (ISO) <https://www.iso.org/fr/home.html>; UNECE ITC WP1 which is the global forum for road traffic safety <https://www.unece.org/trans/areas-of-work/road-traffic-safety/meetings-and-events/global-forum-for-road-traffic-safety-wp1.html>; UNECE ITC WP29 which is the UNECE World Forum for Harmonization of Vehicle Regulations <https://www.unece.org/trans/main/wp29/introduction.html>

2. Rationale

2.1 AI and Criminal Law Responsibility: The Example of Automated (self-driving) vehicles and penal liability

Regrettably thousands of road traffic deaths occur every day all around the world. According to numerous studies human error is the most frequent cause of accidents. In typical vehicular accidents, however, determining which party or parties are at fault is often challenging.

On 19 March 2018 an automated SUV killed a woman in the street in Arizona. She is the first pedestrian known to have been killed by an automated vehicle.[11] The self-driving car was in automated mode at the time of the crash and hit the woman, who was walking across a street outside of the pavement. There was a vehicle operator inside the car at the time of the crash.

As usual in these cases the police must investigate in order to understand the cause of the accident. The first questions are: was the driver driving too fast? was he/she under the influence of alcohol? or drugs? But in the case in Arizona there only was a safety driver. Here we are facing for the first time a situation where an automated vehicle killed a person.

All these incidents raise the same question, a question that's been asked many times before: Who is the responsible party or parties? The car manufacturers? The person/authority who granted the permit to carry out tests? The vehicle operator inside the car at the time of the crash? What is very clear is that for many people there is no clear answer to this question; on the contrary, others consider that it is even easier to identify the responsible party or parties because more data is available.

Testing automated vehicles are clearly operated by the company developing the technology but once the vehicles are bought and owned by individuals the picture of who is to blame becomes even much more unclear. Certain liability issues will be sorted out before automated vehicles are approved for public traffic by the authorities. But the simple question of who is to be held criminally liable for harmful consequences of a machine's autonomous decision-making processes does not always have a simple answer.

2.2 Key Issues in Criminal Law and Artificial Intelligence

As already mentioned, for most technological developments it can be safely assumed that pre-existing criminal law principles and rules will be adequate to ensure liability for serious harm and other forms of unacceptable behaviour. However, this may not be entirely true with respect to Artificial Intelligence, as the autonomous decision-making and self-learning complexity and sophistication at the heart of the technology could leave a responsibility gap.

Across Council of Europe member States, criminal law is generally considered as relating to the conduct, behaviour and intention of human actors, either as natural persons or while acting on behalf of legal persons (corporate liability). This project is concerned with substantive criminal law applicable in all stages of development and utilisation of automated vehicles. The intrinsic

[11] Bryant Smith Walker, 'Uber's Fatal Crash', *Stanford Law School* (19 March 2018), available at: < https://cyberlaw.stanford.edu/blog/2018/03/ubers-fatal-crash>.

complexity of these high-tech systems can lead to significant misunderstandings and misconceptions for many designers, manufacturers, regulators and users, which makes it necessary for all relevant parties to be aware of their respective rights and duties.

Ambiguity and lack of accuracy in these advanced decision-making processes could present major issues of both a factual and legal nature in determining the source of a fault that resulted in harm or damage. While the current generation of smart autonomous robots is capable of a limited degree of autonomous decisions leading to external effects, it is already difficult to conclusively determine the cause of any resulting damage. However, there are considerable challenges presented by the next generation of self-learning robots and self-driving vehicles that could make establishing causation even more difficult.

2.3 Co-operation and co-ordination

Determining appropriate standards for the safe and beneficial use of Artificial Intelligence is a global problem and can only be efficiently countered through increased co-operation and co-ordination, not only between member States but also between the various international organisations and fora involved.

Co-ordinating activities with these and other relevant partners, including the private sector, building on each other's work and avoiding unnecessary duplication, is a clear priority in order for the Council of Europe to add value to the current efforts in these highly complex matters.

3. Stakeholders

Member States have the primary responsibility for ensuring that the many uses of Artificial Intelligence are in compliance with international and national legal standards. It is foreseeable that any process to set regulatory standards in this area will also require input from a range of stakeholders, including, but not limited to:

- Criminal justice system:
    - Prosecutors and investigators,
    - Trial courts,
    - Ministry of Justice / central administrations.
- Education and academia:
    - Robotics engineers,
    - Ethicists,
    - Legal scholars (technology law, information law, criminal lawyers).
- Public authorities:
    - Regulatory agencies,
    - Publicly owned autonomous systems (civil, not military),
    - Government infrastructural systems.
- Private actors:
    - Robotics Manufacturers,
    - Programmers and Software developers,
    - Private companies,
    - AI researchers and development firms.

4. Aim and Objectives

The aim of this project is to determine the principles and rules pertaining to natural and legal persons criminal liability in relation to harm and damage caused by autonomous technologies in a civil context[12], and in particular by automated vehicles.

The objectives of the project are thus to:

1. Examine and ascertain the current existing scope and substance of relevant national criminal legislation and international law pertaining to the use of automated vehicles (or other AI deployment), as well as to determine where and how regulatory powers are established within the competent national public authorities.
2. 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 autonomous technologies, and in what circumstances.
3. Establish where principles and norms of attribution and accountability for natural or legal persons for harm caused by automated vehicles (or other AI deployment) can apply.
4. Examine the scope and substance of an international legal instrument to provide common standards for the criminal law aspects of autonomous technologies and harm caused by artificially intelligent decision-making processes, in particular automated vehicles.

Each of these four main project objectives, activities and the expected outputs/outcomes will be addressed in further detail below.

5. Indicative Logical Intervention

IMPACT – harmonised principles and rules relative to the criminal responsibility for automated vehicles (or other AI deployment) across the Council of Europe area.

5.1 Overall Project Outcome and Outputs

The overall outcome of this project would be to establish an international instrument on criminal offences relating to harm caused in a context of use of Artificial Intelligence and in particular automated vehicles which would be built upon the assessment of the existing international legal framework and national criminal laws of the CoE member states The project is structured along four main outputs:

Output 1

5.1.1 Research project on national criminal law and international legal framework applicable to automated vehicles (or other AI deployment)

Activity: A questionnaire followed by a compilation of responses and analysis.

[12] The reference to a civil context is primarily meant to mean a non-military context: this project does not concern the usage of autonomous functionality by the armed forces of member States.

Reasons: In order to survey the current regulatory framework for Artificial Intelligence, automated machines and in particular automated vehicles, key national-level information should be extracted from the member States.

Working methods: A comprehensive but concise questionnaire is to be developed and distributed to the relevant ministries (or other entities, as appropriate). The results of this questionnaire will be compiled and analysed by an expert or panel of experts.

Expected Output: The final document produced will provide an exhaustive census of relevant national and international legal approaches and instruments, in order to deliver a comprehensive analysis.

Output 2

5.1.2 International Conference on common criminal law standards relating to harm caused by automated vehicles (or other AI deployment)

Activity: Based on the above analysis of relevant national and international legal approaches and instruments, an international conference should be organised providing a forum where member States and also public and private sector actors discuss developments in the field of automated vehicles (or other AI deployment), lacunae in existing criminal law, criminal law solutions already in place and whether there is scope for an international instrument on the criminal law aspects of Artificial Intelligence. Expert input is an essential aspect of the project to ensure that, from an early stage, the project's direction and substantive content is based on the latest and best possible research and knowledge on the subject matter.

Working methods: An international conference bringing together member States and non-member States, private sector actors, and academia.

Expected Output: Conclusions on the need, or not, for the drafting of an international instrument establishing common legal standards in this area.

Output 3

5.1.3 Expert drafting group for an instrument establishing common criminal law standards relating to harm caused by automated vehicles (or other AI deployment)

Activity: Building on the analysis of relevant national and international legislation and the conclusions of the international conference, the Council of Europe could establish an *ad-hoc* drafting group of national experts to develop an international legal instrument providing for appropriate criminal law regulation of the use of automated vehicles (or other AI deployment).

Reasons: Such an international instrument could help ensure a common legal basis for regulatory activity by member States, ensure international co-operation and common criminal law standards between the member States, and also help facilitate mutual legal assistance and international co-operation in criminal matters.

Working methods: Working/Drafting group, composed of representatives of the member States, will meet several times over a defined period.

Expected Output: An international instrument on criminal offences relating to harm caused by Artificial Intelligence and in particular automated vehicles will be prepared.

Output 4

5.1.4 International Conference on the occasion of the adoption of the new international instrument on harm caused by automated vehicles (or other AI deployment)

Activity: International Conference to launch the new instrument, raise awareness of the existence of the instrument and provide explications and information on its provisions and goals.

Working methods: A multi-stakeholder conference bringing together member States and non-member States, private sector actors and academia.

Expected Output: Raised awareness of the new international instrument. States update existing legislation and/or develop new legislation in line with the provisions of the new instrument.