



Strasbourg, 14 September 2018
cdpc/docs 2018/cdpc(2018)14

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 AUTONOMOUS 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

1. Problem analysis and need assessment	3
2. Rationale	6
2.1 AI and Criminal Law Responsibility: The Example of Autonomous (self-driving) vehicles and penal liability	6
2.2 Key Issues in Criminal Law and Artificial Intelligence	6
2.3 Co-operation and co-ordination	7
3. Stakeholders	7
4. Aim and Objectives	8
5. Indicative Logical Intervention	9
5.1 Overall Project Outcome and Outputs	9
Output 1	9
5.1.1 Research project on national criminal law and international legal framework applicable to autonomous vehicles (or other AI deployment)	9
Output 2	9
5.1.2 International Conference on common criminal law standards relating to harm caused by autonomous vehicles (or other AI deployment)	9
Output 3	10
5.1.3 Expert drafting group for an instrument establishing common criminal law standards relating to harm caused by autonomous vehicles (or other AI deployment)	10
Output 4	10
5.1.4 International Conference on the occasion of the adoption of the new international instrument on harm caused by autonomous vehicles (or other AI deployment)	10
5.2 Budget	11

Council of Europe Project on Artificial Intelligence and Criminal Law Responsibility

1. Problem analysis and need assessment

After some incidents with highly-automated driving vehicles in Council of Europe member States¹, and beyond², the following question has been raised: who will be responsible if a completely autonomously driving vehicle³ 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)?⁴

The long-term trends in technological development, and the case of autonomous vehicles, strongly suggest that AI and machines with autonomous functionality will become ever more present in advanced societies, and that States must provide a specific legal framework. A first essential step has already been taken with the introduction of technical standards for special permits allowing automated driving in domestic jurisdictions.⁵ As vehicles quite frequently cross borders, it appears to be in the interest of the member States of the Council of Europe to engage in a standard-setting activity that will ensure their close co-operation in future cases of mutual assistance in criminal matters should autonomous vehicles cause accidents in other countries or illegal activity affect more than one jurisdiction. Unfortunately, it remains highly unlikely that the

¹ 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-milderer-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>.

² 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>

³ According to the SAE standard J3016_201609 < www.sae.org/standards/content/j3016_201609> 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).

⁴ For the purpose of this paper AI may be defined broadly as the attempt to simulate human reasoning. The term has been introduced in a grant application for the famous Dartmouth Conference of 1956: *McCarthy/Minsky/Rochester/Shannon*, 1955 (available at <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>): “We propose that a [...] study of artificial intelligence be carried out [...]. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.”

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>>.

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 offences⁷.

AI and autonomous vehicles are mostly being used in very 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 autonomous 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 unfortunately does not 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 will be a responsibility gap.

Central to this project, and its potential outcomes, is the intent to ensure that regulation is in place to prevent serious harm being caused by autonomous cars or other AI employments without anyone having to answer to the law, and subsequently to avoid undesirable or adverse impacts on the safe use of these advanced technologies.

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 autonomous 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

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>

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.

interests of ensuring adequate means of accountability for situations where autonomous vehicles (or other AI deployment) may cause harm to a human being, it is necessary to help establish a clear criminal-law frame. It is therefore indispensable to set up rules governing any potential criminal liability in advance to ensure that in cases such 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. Duly considering the *ultima ratio* of criminal regulation in this complex field, this project is focused only on the circumstances where the level of harm, or the seriousness of the obligation breached, could or should entail criminal liability.

As the potential widespread adoption of autonomous 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 autonomous 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 autonomous 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 autonomous 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 autonomous vehicles.

The technology standards developed at international level¹⁰ could then pave the way for suitably careful legal regulation at national level relating to the use and oversight of autonomous driving including principles for allocating criminal law responsibility for the employment of AI and appropriate sanctions and measures, where necessary, as well as

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](http://www.europarl.europa.eu/RegData/etudes/STUD/2018/615635/EPRS_STU(2018)615635)>, p. 42.

a common legal framework to resolve any transborder issues and a legal basis to facilitate mutual legal assistance in criminal matters.

2. Rationale

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

Regrettably thousands of car accidents happen 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 is at fault is often challenging. This challenge grows when more drivers rely on semi-autonomous driving and (some) vehicles are driven by AI.

On 19 March 2018 an autonomous SUV killed a woman in the street in Arizona. She is the first pedestrian known to have been killed by an autonomous vehicle. The self-driving car was in autonomous 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 was no driver. Here we are facing for the first time a situation where a self-driving car killed a person.

All these incidents raise the same question, a question that's been asked many times before: Who is the responsible party? 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 there is no clear answer to this question.

Testing autonomous 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 autonomous 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 unfortunately does not have a simple answer. It is certainly difficult to deal with "criminal behaviour" of "non-human beings".

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 leaves a responsibility gap. Or phrased differently: if we lose the human in the driver seat, who will be responsible for the driving? This responsibility gap may necessitate the creation of a supplementary *sui generis* legal regime.

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 autonomous vehicles. The intrinsic 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:
 - o Prosecutors and investigators,
 - o Law enforcement agencies,
 - o Judicial entities.
- Education and academia:
 - o Robotics engineers,
 - o Ethicists,
 - o Legal scholars (technology law, information law, criminal lawyers).
- Public authorities:
 - o Regulatory agencies,
 - o Publicly owned autonomous systems (civil, not military),
 - o Government infrastructural systems.
- Private actors:
 - o Robotics Manufacturers,
 - o Programmers and Software developers,
 - o Private companies,
 - o AI researchers and development firms.

4. Aim and Objectives

The aim of this project is to determine the principles and rules pertaining to individual criminal liability in relation to harm and damage caused by autonomous technologies in a civil context¹¹, and in particular by self-driving 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 autonomous 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 autonomous 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

¹¹ 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.

harm caused by artificially intelligent decision-making processes, in particular autonomous 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 autonomous 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 by artificial intelligence and in particular autonomous vehicles is 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 autonomous vehicles (or other AI deployment)

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

Reasons: In order to survey the current regulatory framework for artificial intelligence, autonomous machines and in particular self-driving 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 a comprehensive analysis of relevant national and international legal approaches and instruments.

Output 2

5.1.2 International Conference on common criminal law standards relating to harm caused by autonomous 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 autonomous 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 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 autonomous 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 expert drafting group to develop an international legal instrument providing for appropriate criminal law regulation of the use of autonomous 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 autonomous 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 autonomous vehicles (or other AI deployment)

Activity: 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 multistakeholder 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.

5.2 Budget

2 International Conferences (output 2 and output 4):

110 000€ x 2 = 220 000€

4 Expert Drafting Group meetings (output 3): 50 000€ x 4 = 200 000€

1 Research Project (output 1): 10 000€ x 1 = 10 000€

Translations/Documentation: 5 000€ x 1 = 10 000€

1 Project Officer (grade B5): 24 months = 160 800€

1 Administrative Assistant (grade B2): 24 months = 112 800€

Total= 713 600€

7% administrative costs: 49 952€

Grand Total = 763 552€