þÿ A u t o m a t e d v e h i c l e s i s a d i l u t i o n o f h u m a n r e s p

Günsberg, Patrick

2022

þÿ G ü n s b e r g , P 2 0 2 2 , ' A u t o m a t e d v e h i c l e s i s a d i l u t i o n o f h u m a n r e s p
þÿ ' , N e w J o u r n a l o f E u r o p e a n C r i m i n a l L a w , v o l . 1 3 , n o . 4 , p p . 4 3 9 4 5

http://hdl.handle.net/10138/351453
https://doi.org/10.1177/20322844221138049

Article

NJECL

New Journal of European Criminal Law
2022, Vol. 13(4) 439–451


Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20322844221138049
journals.sagepub.com/home/nje
SAGE

Automated vehicles – is a dilution of human responsibility the answer?

Patrick S. Günsberg
University of Helsinki, Finland

Abstract

AI-driven vehicles and other artificial intelligence (AI) systems may cause serious injury to people while operating independently. Besides vehicles progress may be seen in the use of autonomous weapon systems, AI in medicine and care robots. It seems that soon AI systems will increasingly be making decisions previously made by humans. A Swedish inquiry argued that existing criminal law rules on responsibility are not suitable for automated vehicles (when in the self-driving mode). The human in the driver's seat would not be blamed if an accident occurs. Conversely, the Proposal for a Regulation on Artificial Intelligence places an emphasis on oversight by human beings to an extent. A battle for the hearts and minds of people might be underway here. It seems that further exploration of the matter is warranted, especially through the criminal law lens—are proposals where the human user is absolved of blame viable at this point in time?

Keywords

Artificial intelligence, criminal responsibility, dilution of human responsibility, negligence, strict liability

Corresponding author:
Patrick S. Günsberg,
Email: patrick.gunsberg@helsinki.fi

Introduction

AI-driven vehicles and other artificial intelligence (AI) systems may cause serious injury to people while operating independently.[1] Besides vehicles progress may be seen in the use of autonomous weapon systems, AI in medicine and care robots. It seems that soon AI systems will increasingly be making decisions previously made by humans.[2] A Swedish inquiry argued that existing criminal law rules on responsibility are not suitable for automated vehicles (when in the self-driving mode).[3] The human in the driver's seat would not be blamed if an accident occurs.[4] Conversely, the Proposal for a Regulation on Artificial Intelligence places an emphasis on oversight by human beings to an extent. A battle for the hearts and minds of people might be underway here. It seems that further exploration of the matter is warranted, especially through the criminal law lens—are proposals where the human user is absolved of blame viable at this point in time?[5]

Structure and methodology

The paper first considers the responsibility of the users of automated vehicles that operate independently. Particularly, the proposals to not to hold the human in the driver's seat responsible for traffic violations will be evaluated. Second, the viability of strict liability in this context will be addressed—should a natural person be held responsible for an unfortunate outcome? Finally, the paper investigates the balance between acceptable and unacceptable risks—does it seem that our societies would be ready to accept that the human in the driver's seat should no longer be held responsible? A comparatively oriented analysis will be used to shed light on the aforementioned

1. Regarding definitions of AI, see Russel Stuart and Norvig Peter, *Artificial Intelligence A Modern Approach,* Third Edition Pearson Education Limited 2016, pp. 1-5; Chesterman Simon, 'Artificial Intelligence and the Limits of Legal Personality', *International and Comparative Law Quarterly, Vol. 69,* 2020, p. 1; Simmler Monika and Markwalder Nora, 'Guilty Robots? – Rethinking the Nature of Culpability and Legal Personhood in an Age of Artificial Intelligence,' *Criminal Law Forum* 30: 1-31, 2019*,* p. 9; Pasquale Frank, 'Toward a Fourth Law of Robotics: Preserving Attribution, Responsibility, and Explainability in an Algorithmic Society, *Ohio State Law Journal,* Vol. 78:5, 2017, p. 1254; Gal, Michal S., 'Algorithmic-Facilitated Coordination: Market and Legal Solutions', *CPI Antitrust Chronicle May,* 2017; See Ripatti Mikko, 'Künstliche Intelligenz: Algorithmen für den Aktienmarkt', October 10, 2018, available at https://www.wallstreet-online.de/nachricht/10917930-kuenstliche-intelligenz-algorithmen-aktienmarkt, last visited on August 19, 2022.
2. Beck Susanne, 'Die Diffusion strafrechtlicher Verantwortlichkeit durch Digitalisierung und Lernende Systeme', *Zeitschrift für Internationale Strafrechtsdogmatik,* 2020, p. 41.
3. Ds 2021:28, p. 150; Automated vehicles may also be called autonomous vehicles and driverless vehicles, but for instance the UK joint report argued that the term 'autonomous' could be viewed as too anthropomorphic, see Law Commission and Scottish Law Commission, *Automated Vehicles: Joint report,* 25 January 2022, p. 1, 18; Chesterman Simon, 'Artificial Intelligence and the Problem of Autonomy,' *Nortre Dame Journal on Emerging Technologies* Vol. 1, 2020*,* p. 3; SOU 2018:16, p. 70.
4. SOU 2018:16, p. 95.
5. Potential alternatives to human responsibility, such as the criminal liability of AI systems are beyond the scope of this paper. It seems natural to first investigate how such a dilution of human responsibility would be received in our societies; See the Proposal for a Regulation of the European Parliament and of the Council laying down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM(2021) 206 final, Brussels, 21.4.2021; For instance, the Council of Europe has launched an investigation into the said matter which reflects the need for a more thorough understanding of the difficulties involved, see Council of Europe, *European Committee on Crime Problems, Questionnaire concerning Artificial Intelligence and Criminal Justice (using the example of Automated Driving),* CDPC(2019)8FIN Strasbourg, 19 May 2019; See also European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM(2021) 206 final, Brussels, 21.4.2021.

issues. Inter alia, the United Kingdom and Swedish investigations into automated vehicles will be drawn upon.[6] Jurisdictions that offer fertile ground for a comparison have adequately different legal solutions. In terms of the United Kingdom and Sweden, the former belongs to the common law legal family and the latter to the civil law one. This contrast has the potential to enable a fruitful discussion of possible legal solutions.[7] It should be noted that comparative research aims to make functional comparisons—in other words laws with comparable purposes may be usefully compared. This premise sets limits for instance to the case law that is subject to a comparison in the present analysis.[8]

Automated vehicles and the reach of criminal law

The wider society has a stake in holding the person responsible who controls something that constitutes a threat. Therefore when it comes to negligence a human centered criminal liability seems a good starting point in terms of AI systems.[9] It should be noted, however, that a possible design of criminal responsibility for humans would have to take note of factors that are not constant—for instance, the variety of AI systems in terms of their stage of development and the power that humans have over them are of relevance.[10]

Potential future criminal regulation could make the reckless utilization of AI systems a criminal offence. Such regulation would, however, aim at blameworthy individual behavior, but would fail to address a situation where AI systems act independently. One possibility is that a proximate person who would normally escape criminal liability, would be responsible. For those employing and designing AI systems, it could be made mandatory to designate a person for this purpose.[11] If due to the number of developers in the creation of an AI system one designer cannot be designated, one could for instance hold the user liable.[12]

It is still unclear, however, who will exercise control over decision-making in terms of automated vehicles in the future, but it will have an impact on the design of criminal liability. The outcome may vary from one manufacturer to the next. For example, in the aviation industry Boeing and Airbus

6. See Günsberg Patrick, *Criminalizing Business Cartels in Europe – A Comparative Perspective,* Helsinki: Unigrafia 2015, pp. 4ff.; See the discussion in Eser Albin, 'The Importance of Comparative Legal Research for the Development of Criminal Sciences', *Nouvelles Études Pénal,* 1998 and Zweigert K & Kötz H. *An Introduction to Comparative Law.* New York: Oxford University Press, 3rd, ed. 1998; See Regeringskansliet, 'Ansvarsfrågan vid automatiserad körning samt nya regler i syfte att främja en ökad anvädning av geostaket', *Ds 2021:28;* Law Commissions, 2022.
7. Günsberg 2015, p. 5.
8. The solutions of different jurisdictions should be explored without the strain of context and doctrine. Rather, a comparative analysis should exclusively focus on the designated function of the discovered legal solutions Günsberg 2015, p. 8; Zweigert & Kötz 1998, p. 44.
9. Fateh-Moghadam Bijan 'Innovationsverantwortung im Strafrecht: Zwischen strict liability, Fahrlässigkeit und erlaubtem Risiko – Zugleich ein Beitrag zur Digitalisierung des Strafrechts', *ZSTW* 2019, pp. 886-887.
10. Lima Dafni, 'Could AI Agents Be Held Criminally Liable: Artificial Intelligence and the Challenges for Criminal Law,' *South Carolina Law Review,* Vol 69:677, 2018, p. 693; See also Sabine et al., 'If Robots Cause Harm, Who is to Blame: Self-Driving Cars and Criminal Liability', *New Criminal Law* Review, 2016, p. 433.
11. Abbott, Ryan Benjamin and Sarch, Alex F., Punishing Artificial Intelligence: Legal Fiction or Science Fiction, *University of California, Davis,* Vol. 53:323, 2019, p. 378; European Commission Proposal for a Regulation on Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), Brussels, 21.4.2021, COM(2021) 2016 final; Pasquale 2017, pp. 1253-1254.
12. See the discussion in Günsberg, Patrick, Some Thoughts on Artificial Intelligence (AI) Systems and Responsibility (February 2, 2019). Available at SSRN: https://ssrn.com/abstract=3329176 or http://dx.doi.org/10.2139/ssrn.3329176, last visited on August 19, 2022; Abbott and Sarch 2019, p. 378.

diverge on the question who controls the aircraft. Boeing entrusts human beings with more responsibilities, while Airbus, places greater trust on the autopilot.[13] It seems that the period preceding fully automated vehicles will see joint human-machine control. This means a shift from the prevailing order where humans are in exclusive control of their vehicles to one where machines and humans may split the responsibility.[14]

Conditional automation and the role of the human in the driver's seat

The Society of Automotive Engineers International (SAE) has introduced general terminology for a discussion on driving automation. This terminology takes the shape of a classification of six levels of driving automation. The classification is helpful, inter alia, for a comparative analysis.[15] A decisive line is drawn between features that assist the driver (levels 0–2) and automated driving (levels 3–5).[16] Technologies that assist driving concern for instance cruise control, automatic braking and lane keeping systems. Such technologies do not replace the driver. These support features must be continuously supervised by the human driver.[17] With SAE level 3 or "conditional automation" regulators face a conundrum—the vehicle can drive itself but requires a human in the driver's seat to fall back on. Under the SAE classification the human in the driver's seat would not be required to oversee the driving environment.[18] In contrast, SAE level 3 could also be regarded as a type of driver assistance which requires human oversight (as various levels of automation exist on a continuum). That would mean that the human in the driver's seat would be required to constantly keep an eye on the surroundings. For instance, text messaging would be banned. Thus, under criminal law drivers' duties would remain unchanged.[19] The reasoning in that case would be that if a vehicle needs a human to fall back on it is not actually automated and should not be considered to "safely drive itself."[20] For instance, Volvo cars, a consultee, argued that a system is not an automated one if the human driver needs to sort out "conflict situations" when the self-driving mode is engaged.[21] The Swedish discussion noted that potentially one of the most important questions in the whole matter lies in the role of the driver—what will the human in the driver's seat be held responsible for?[22]

The UK consultation paper suggested that the user-in-charge (=the human in the driver's seat)[23] would not be responsible for violating driving rules when the self-driving mode is engaged. If it seems that the automated driving system is defective a regulatory authority would take over the case.

13. SOU 2018:16, pp. 542-543.
14. Awad Edmond et al., 'Drivers are blamed more than their automated cars when both make mistakes', *Nature Human behaviour*, Vol. 4 February 2020, p. 134.
15. See Society for Automotive Engineers International (SAE), J3016 Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, April 2021; Law Commissions 2022, p. 11.
16. Law Commissions 2022, p. 13.
17. Ds 2021:28, p. 152; Law Commissions 2022, p. 12.
18. From a safety perspective for level 4 vehicles, it is not equally serious (as it is in terms of level 3) if the user fails to take over driving tasks when requested by the driving system, see Law Commissions 2022, p. 13.
19. See Law Commissions, 'Automated Vehicles: Analysis of Responses to the Preliminary Consultation Paper', 19 June 2019, p. 32; Law Commissions 2022, p. 35.
20. Law Commissions 2019, p. 34.
21. Ibid., pp. 35-36.
22. Ds 2021:28, p. 168.
23. The notion of a 'user-in-charge' concerns the human in the driver's seat at a certain phase of automated driving. That phase sets the human free from monitoring the surroundings and does not require undertaking driving tasks with little warning. The human needs to be prepared for a planned handover, however, see Law Commissions 2019, p. 37.

The regulatory authority would have a range of sanctions at its disposal that could target the automated driving system entity (ADSE) behind the automated driving system.[24] Similarly, the Swedish inquiry (SOU 2018:16) suggested that humans should be either outside or inside the decision loop.[25] Thus, the human in the driver's seat would cease to be responsible when the self-driving mode is engaged.[26] The UK report's reasoning was, inter alia, that SAE level 3 systems do not require much from humans in the driver's seat which may leave them daydreaming, for example. Therefore, it was argued that secondary activities (during automated driving) could be "a way of managing drivers' attention." It was noted that research indicates that for people monitoring a task is more problematic when done passively than actively. Arguably, human attention is either turned on or off.[27] However, for example, The Transport Safety Research Group, a consultee, noted that there is compelling evidence which indicates that secondary activities prevent concentration.[28] Also, a slight majority of 52% of the consultees felt that "there should be no relaxation of the laws against distracted driving for systems which relied on human intervention to be safe."[29] It could be argued then that the UK report's reasons for the profound changes regarding the role of the human in the driver's seat are not well justified.

Indeed, the Faculty Committee of the University of Stockholm, a consultee, argued that the driver should have a continuous oversight duty that aims to prevent traffic offences and accidents. Arguably, the importance of this is underlined by the developmental phase where the technology is new and untested. The risk of defects and accidents that are related to automated driving is relatively high.[30] For this reason too, proposals which would not require the user to monitor the driving environment seem premature. Moreover, the idea that no one would be responsible for how the vehicle travels when the self-driving mode is engaged does not appear a possible option because such an arrangement would not contribute to traffic safety and could open the door to abuse.[31] An example of abuse could be where the human user would let the vehicle drive too fast while not being held responsible.[32] This danger appears to further erode the sensibility of the aforementioned UK proposal.

24. Ibid. 2019, p. 107.
25. The human decision-making power depends on the given technology. In terms of military technology three levels exist, which may offer a reference point here. Firstly, human beings may exercise full control of what actions a robot takes (human in the loop). Secondly, there is the possibility that the robot has the ability to independently choose its targets without human interference, but the human would retain the possibility to watch over and may interfere with steering (human on the loop). Thirdly, humans can be left outside decision making, which means that robots take action without human interference, see SOU 2018:16, p. 542, 640; see also https://machine-rockstars.com/lexikon/was-ist-human-in-the-loop/, last visited on August 19, 2022; Beck 2020, p. 49.
26. SOU 2018:16, p. 95.
27. Law Commissions 2019, pp. 32ff; Law Commissions 2022, pp. 35-36, 38ff.
28. Law Commissions 2019, pp. 36-37.
29. Law Commissions 2019, p. 13.
30. Stockholms universitet, 'Remiss: Vägen till självkörande fordon-introduktion (SOU 2018:16), August 24, 2018.
31. It may not be ruled out that people adapt their driving so that if an accident takes place the machine will be blamed – an example of that would be if people refrained from trying to remedy mistakes that resulted from an override by the machine. Possibly such a development is already taking place in judicial decision-making. As judges decide whether to set people free or keep them locked up while they await trial, the judges often depend on Actuarial Risk Assessment Tables in their decision making. It has been suspected that judges unduly rely on the aforementioned tables in an effort to avoid responsibility if the person once set free engages in criminal activity, see NPR, 'Did A Bail Reform Algorithm Contribute To This San Francisco Man's Murder?', August 18, 2017; Awad Edmond et al., 2020, pp. 139-140.
32. It may be noted that Zurich Insurance (UK) argued that muddy sensors and obsolete software should make the user criminally responsible even if self-driving mode is engaged, see Law Commissions 2019, pp. 111-112.

AI and strict liability

As AI systems are increasingly autonomous it becomes more difficult to predict how they act. For instance, an automated vehicle is in contact with its environment and employs the data available to it and will act independently on it. Simultaneously, questions regarding proper standards of care by manufacturers, users and programmers become more pertinent. It could be argued that a responsibility gap emerges as autonomous decision-making systems make attributing responsibility to a human being for wrongdoing more complicated. Even discovering any wrongdoing might be a problem.[33]

It is interesting then that the Swedish inquiry built on the strict liability of the vehicle owner (though not under criminal law in name at least) if the vehicle does not comply with traffic rules when the self-driving mode is engaged. Similarly, the UK investigation sought to rely on regulatory sanctions. The UK joint report aimed to encourage "a no-blame safety culture" which draws lessons from errors. The idea was that the fact that a human driver would be criminally prosecuted in a given situation does not mean that the ADSE should also be subject to blame.[34] In the Swedish memorandum's view, however, once the producer is no longer in control of the vehicle and how it is used the vehicle owner is more appropriately to be held responsible for potential traffic violations.[35] The Swedish solution seems better in the sense that automated vehicles may evolve to become very different from what their creators first envisioned due to the exposure to new influences.[36]

Yet, several responses to the consultation exercise pointed out that it does not appear reasonable to hold the vehicle owner liable when the error which caused the vehicle to violate a traffic rule is based on how the producer programmed the driving system.[37]

Indeed, the subsequent Swedish memorandum sought to modify the inquiry's proposal by removing the vehicle owner's need to pay a penalty fee when the violation of traffic rules is based on an error in the automated driving system which is beyond the owner's control. The memorandum underlined the notion of an administrative penalty fee as an economic sanction, while their penal nature under the case law of the ECtHR was acknowledged.[38] It has been noted that a shift from criminal penalties to administrative penalties may be problematic, however, as robust criminal law protections would be absent while administrative sanctions could be harsh.[39] This casts a shadow of doubt on the Swedish proposal as do potential evidential problems—evidential problems may occur

33. Gless Sabine et al., 'If Robots Cause Harm, Who is to Blame: Self-Driving Cars and Criminal Liability', *New Criminal Law* Review, 2016, p. 426; Fateh-Moghadam Bijan 2019, p. 876; Williams Rebecca, 'Lords Select Committee, Artificial Intelligence Committee, Written evidence, 2017, available at http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/artificial-intelligence-committee/artificial-intelligence/written/70496.html#_ftnref20, last visited on August 19, 2022.
34. Ds 2021:28, p. 235; see Law Commissions 2022, pp. 209-210.
35. Ds 2021:28, pp.229-230.
36. Pasquale 2017, p. 1254.
37. In the proposal (SOU 2018:16) there was a possibility for adjustment – in other words the administrative penalty fee would not be collected if it is unreasonable (5 kap. 12§). The availability for adjustment and judicial review in a sanction system that builds on strict liability may be regarded as a requirement of compliance with the case law of the European Court of Human Rights on the presumption of innocence and the right to a fair trial. The subsequent Swedish memorandum took the position, however, that it would not be a satisfying outcome if the vehicle owner would have to rely on an adjustment taking place or would be required to initiate judicial proceedings, Ds 2021:28, pp. 235-236.
38. Ibid, pp. 227-228, 235, 236.
39. Lahti Raimo, 'The origin and development of quasi-criminal law enforcement mechanisms in Europe: Nordic perspective, in Franssen and Harding ed. *Criminal and Quasi-Criminal Enforcement Mechanisms in Europe: Origins, Concepts, Future,* Bloomsbury Publishing Plc: Oxford 2022, p. 23; Günsberg 2015, p. 97.

when an attempt is made to show that a traffic violation was beyond the vehicle owner's control because as was noted even detecting any wrongdoing could be demanding.

Moreover, it would be dubious to shift from driver's responsibility to vehicle owner's liability. Arguably, the driver should be required to intervene when faced with dangers even if the vehicle does not require it.[40] Also, the decision to use an automated vehicle and creating risks for others should make the human user responsible for adverse implications.[41]

Strict liability in relation to AI could be justified by the claim that there is an obligation to the public to come up with guarantees that particularly severe risks are reduced as far as is feasible. It has been put forward that in an interactional society we are under an obligation to guarantee others that we behave in a way that is not likely to cause harm.[42] Abbott and Sarch observed the possibility that a designated person, "a Responsible Person," that is held strictly criminally liable could provide those guarantees.[43] Thus, it may be helpful to discuss what strict liability could mean in this context.

De facto strict liability

The Swedish inquiry argued that if humans were on the decision loop it would make such heavy demands on the human in the driver's seat that it would border on strict liability when the self-driving mode is engaged. Arguably, this would mean criminalizing unconscious negligence and would conflict with the principle of culpability.[44]

Interestingly, the mildest forms of strict liability come close to unconscious negligence.[45] When an actor is unconsciously negligent, he is not at all aware that he is violating a norm of due care.[46] For instance, a driver could exceed the speed limit unconsciously in traffic while he would not violate the relevant rules consciously. Thus, the actor is blamed because he was negligent despite the ability to comply with the standards of due care. Unconscious negligence does not necessarily attract less blame than conscious negligence. A driver who is indifferent to possibly posing a danger may create a more serious risk to the safety of others than a driver who evaluates risks on a continuous basis.[47] If the

40. Ds 2021:28, p. 226.
41. Law Commissions 2019, pp. 111-112; Incidentally, similarly in terms of autonomous weapons, it has been noted that their use involves an intrinsic risk that even utmost care will not eliminate. It appears that since the possibility of harm can be reduced solely by the one who initiated the action, it is appropriate that the same entity bears the responsibility for the outcome, see Crootof 2016, pp. 1395-1396; See also the discussion in Günsberg, Patrick, Some Thoughts on Artificial Intelligence (AI) Systems and Responsibility (February 2, 2019). Available at SSRN: https://ssrn.com/abstract=3329176 or http://dx.doi.org/10.2139/ssrn.3329176.
42. Abbott and Sarch 2019, p. 380; Anthony Duff discussed this idea in terms of mala prohibita offences, see Duff Anthony, *Answering for Crime*, Hart Publishing, Portland, 2007, p. 170; Tapani Jussi, Tolvanen Matti and Hyttinen Tatu, *Rikosoikeuden yleinen osa – Vastuuoppi*, Alma Talent, Helsinki 2019, p. 332; Kaiafa-Gbandi, 'Artificial intelligence as a challenge for Criminal Law', in ed. Beck Susanne et al., *Digitalisierung, Automatisierung, KI und Recht*, Nomos Verlagsgesellschaft, Baden-Baden 2020, p. 323; It may be noted that on a global scale strict criminal responsibility for serious offences is not completely absent. For instance, a court ruling in Alabama noted that the legislation which prohibits selling alcohol to children reflects the intention of the legislator to employ strict liability. Statutory rape has also been mentioned as an example, see Abbott and Sarch, 2019 p. 380; Funari v. City of Decatur, 563 So. 2d, 54, Court of Criminal Appeals of Alabama.
43. Abbott and Sarch, 2019, pp. 380-381; See also Simons Kenneth W., 'When is Strict Criminal Liability Just?', *Journal of Criminal Law and Criminology,* Vol 87, Issue 4 Summer 1997, pp. 1075-1076.
44. SOU 2018:16, p. 640; Ds 2021:28, p. 224, 226, 227.
45. Tapani et al. 2019, p. 339.
46. Ibid. 2019, pp. 331-332.
47. Formulating the principle of trust would be very difficult if for instance only a conscious violation of a duty to give way was punished, Tapani et al. 2019, p. 332; Regarding the principle of trust see also Kaiafa-Gbandi, 2020, p. 323.

criminal offence required that the conscious action by the actor was proved, a significant part of even serious traffic violations might go unpunished.[48] In this sense the argument that a human user's oversight responsibility would be unreasonable is not convincing.

In the common law context, it has been noted that when demonstrating mens rea is demanding, one may resort to strict liability—that would mean that criminal intent would not need to be established.[49] By way of example the House of Lords considered in *Empress Car* that the appellants were responsible for oil entering a river despite the leakage being brought about by the sabotage of a third party.[50] While that approach has been subject to criticism by criminal law scholars, it could possibly offer a blueprint for deciding cases involving AI systems.[51] Yet in the context of criminal law the principle of culpability is frequently invoked to reject the introduction of strict liability.[52] Strict liability offences are a type of criminal offence that can be found in the Anglo-American legal systems whereas they are absent in the Nordic context.[53]

In *Empress Car* Lord Hoffmann held that since the firm kept up a diesel container it amounted to "doing something" which justified conviction.[54] Lord Hoffmann argued that the intentional act of the third person which caused the contamination does not suggest that the defendant, having created the circumstances where the third person could bring about the contamination, was not causing the contamination as well (in the sense of the relevant rules). As opposed to "absolute liability" it is not adequate to only demonstrate that leakage took place, regardless of the way it came about. One should be able to maintain that the contamination was caused by the defendant. Lord Hoffmann pointed out that the matter boiled down to the defendant having caused the contamination and not whether the defendant should have foreseen such a result.[55]

Negligence offences and the way of dealing with the subject matter in *Empress* aim to tackle the executives' omission to take adequate preventive measures. Arguably, when an individual undertakes a particular activity, he should reckon that certain responsibilities follow.[56] That approach seems even more justified if the risky activity greatly benefits the person behind the activity.[57] One could turn the omission into a criminal offence. In terms of the omission, for instance the omission to apply reasonable consideration to avoid causing harm, the difficulty is that no matter to what lengths the defendant goes in that regard it may often fail to prevent the harm. That approach would also risk

48. This then could have an influence on how people view the importance of complying with the relevant rules and the ability of the criminal justice system to maintain compliance with norms, Tapani et al. 2019, p. 332.
49. Williams 2017; See also Fateh-Moghadam Bijan 2019, pp. 885-886.
50. *Environmental Agency (formerly National Rivers Authority) v Empress Car Co (Abertillery) Ltd [1999]*; Williams 2017.
51. Williams 2017; See Ormerod David and Laird Karl, *Smith and Hogan's Criminal Law,* 14th ed. Oxford University Press: Oxford 2015, p. 100.
52. Fateh-Moghadam Bijan 2019, pp. 881-882.
53. Tapani et al. 2019, p. 340.
54. It was required, however, that causal link was shown to exist. Lord Hoffmann indicated that the section 85(1) of the Water Resources Act 1991 imposes strict liability and therefore does not set out mens rea requirements either in the form of deliberateness or negligence.
55. *Environmental Agency (formerly National Rivers Authority) v Empress Car Co (Abertillery) Ltd [1999]*; Williams 2017; The House of Lords indicated subsequently that the analysis in *Empress* applies to cases involving pollution, but is not applied generally across criminal law. Thus, the principle in *Empress* was not applicable in cases concerning manslaughter, see Ormerod David and Laird Karl, *Smith and Hogan's Criminal Law,* 14th ed. Oxford University Press: Oxford 2015, p. 100.
56. Ashworth Andrew, *Positive Obligations in Criminal Law,* Hart Publishing: Oxford 2013, p. 79; Williams 2017.
57. Williams 2017; See also Fisse Brent, 'Reconstructing Corporate Criminal Law: Deterrence, Retribution, Fault, and Sanctions,' *Southern California Law Review, Vol 56:1141* 1983, pp. 1203-1204.

firms merely performing such activity perfunctorily rather than genuinely seeking to avoid the harm. As an alternative one could make it a criminal offence that the defendant did not keep the harm from happening.[58]

Interestingly, the Finnish Supreme Court discussed the requirement of foreseeability in terms of responsibility for negligent homicide in one case.[59] It was considered whether the piling up of snow and ice and ice falling on a person while visiting a building with fatal results was an unexpected incident. The court took the position that the incident was not unforeseeable.[60] The chairperson of the housing company that was singled-out as the responsible person. He had a legal duty to prevent the result.[61] The court argued that the central place of the building was liable to underline the safety duty. Also, the weather at the time was to be taken into account.[62] The standards set could be viewed as burdensome—the defendant did not even live in the same town where the building was situated.[63] Similarly, the difficulty of foreseeing how an automated vehicle behaves does not mean that, for example, users, designers, or owners should not be held accountable since the inability to predict the behavior of automated vehicles means precisely that one should exercise due care.[64] Under this reasoning, one could argue that since users, designers and owners are aware the automated vehicle could cause harm they could be held accountable—arguably this constitutes "de facto strict liability."[65]

In terms of driverless vehicles, the Faculty Committee of the University of Stockholm raised the question of the driver's duty to react when faced with a potential accident (under the general part of criminal law). Could failing to do so have the driver prosecuted for manslaughter? Clearly, until now the driver has held a position of responsibility.[66]

Subsequently, the Swedish memorandum acknowledged the human in the driver's seat is considered to have a position of supervision responsibility. This idea underlies the role of a driver and also the human in the driver's seat is covered by the basic rules for road-users in the traffic regulation. Among these rules may be found the general rule of traffic which sets out that a road-user must observe the level of care that the circumstances require to avoid traffic accidents.[67] Interestingly, as the relevant case law has it a road user may be found not guilty of the offence of negligence in traffic while still be guilty of negligent homicide or bodily injury caused through negligence. In terms of the latter offences, it is not required that the involved negligence is of the

58. Williams pointed out that a defence of impossibility should, however, be available. In other words, the defendant should not be held responsible for harm that he could not have kept from happening, Williams 2017; See also Fisse Brent, 'Reconstructing Corporate Criminal Law: Deterrence, Retribution, Fault, and Sanctions,' *Southern California Law Review, Vol 56:1141* 1983, pp. 1203-1204.
59. It may be noted that Nordic omission offence tenets draw to a great extent on the German criminal law doctrine, see Nuutila Ari-Matti, Rikosoikeudellinen huolimattomuus, Lakimiesliiton kustannus 1996, p. 232ff.
60. Lahti Raimo, 'Asunto-osakeyhtiön hallituksen puheenjohtaja kuolemantuottamusvastuussa turvallisuusvelvoitteen laiminlyönnistä', in Timonen Pekka (ed.) *KKO:n ratkaisut kommentein 2007,* Lakimiesliiton kustannus 2007.
61. Lahti Raimo, 2007; KKO 2007:62, at 5 and 7.
62. Lahti Raimo, 2007; KKO 2007:62 at 8 and 20.
63. It may be noted that the principle of last resort (ultima ratio) indicates that the extent of responsibility for negligence in criminal law ought to be more limited than in tort law, Lahti Raimo, 2007; Regarding 'the ancillary nature of criminal law' see also, Kaiafa-Gbandi, 2020, p. 319.
64. Gless Sabine et al. 2016, p. 427.
65. Ibid, p. 427.
66. Stockholms universitet, 'Remiss: Vägen till självkörande fordon-introduktion (SOU 2018:16), August 24, 2018; See 3 kap. Section 7 and 8 of the Swedish Penal Code.
67. Ds 2021:28, p. 218.

aggravated type although not every little misjudgment can be considered negligent in the way that criminal responsibility is required.[68] The Swedish memorandum also noted that the human in the driver's seat might not be absolved of blame if he or she is aware that the vehicle is plagued by a safety defect that creates a risk for the violation of traffic rules.[69] Furthermore, the car maker could be aware that the automated driving system is plagued by a serious safety defect but chooses to ignore the risk and makes the vehicle available to buyers. The rules of the Penal Code could be applicable in an accident where someone is injured.[70] Yet the Swedish memorandum (Ds 2021:28) underlined that as a starting point the human in the driver's seat would not be responsible for driving when the automated mode is engaged and could engage in secondary activities.[71]

There seems to be an ongoing strain between arguments advocating strict liability on consequentialist grounds and those underlining the importance of demonstrating culpability. One often adduced objective with strict liability is to ensure public protection.[72] It seems reasonable to assume that reasoning similar to the one cited in the Finnish case could be appropriate in terms of the use of automated vehicles and should not be rejected out of hand.[73]

Acceptable risk and automated vehicles

Life is fraught with risks. People also expose other people constantly to various hazards. Criminalizing a given conduct is all about striking a balance between the risks that the society is willing to take and unacceptable risks. In our daily lives we take risks because we expect the utility of a given technology to exceed the potential risks for harm. For instance, using a car has its advantages as it entails speedier transportation. Yet the disadvantages include potential traffic accidents.[74] It is not considered immoral or criminal to manufacture or use vehicles even though they become involved in accidents with lethal consequences—this owes to the fact that the risks involved are accepted by the society.[75]

Yet we do not know how things unfold as automated vehicles start driving on public roads on a large scale. Diverging views exist regarding the peril posed by automated vehicles.[76] Measuring a vehicle's safety in terms of a specified standard prior to an authorization is difficult.[77]

The general public as a whole has an interest in technological advancement, but also values products that are safe. As a result, it could seem reasonable to argue for a compromise in terms of criminal liability in cases involving negligence.[78] For that reason, it may be justifiable to argue in favor of putting a limit to for instance the producers or owners' liability under criminal law. That

68. Ibid, p. 203.
69. Ibid, p. 205.
70. Ibid, p. 200; SOU 2018: 16, p. 700, 703.
71. Ds 2021:28, p. 219.
72. Ashworth Andrew, 'Should Strict Criminal Liability be Removed from All Imprisonable Offences?,' Vol. *45 Irish Jurist,* 2010, p. 9.
73. Kaiafa-Gbandi argued, however, that her 'analysis upholds that harmful outcomes of AI systems' use in modern technological societies could remain unaddressed by means of ascribing criminal liability on any of the persons involved in the sequence which begins with their programming or production and ends with their use, especially if one seeks to preserve the fundamental principles of a criminal law that abides by the rule of law,' see Kaiafa-Gbandi, 2020, p. 323.
74. SOU 2018:16, p. 550.
75. Ibid, p. 551.
76. Ibid., p. 551; Ds 2021:28, pp. 231-232.
77. Law Commissions 2022, p. 56.
78. The inability to predict the course of action a smart machine will take, means that potential harmful conduct by the same machine remains foreseeable, see the discussion in Gless et al., 2016.

objective could materialize if duties of care were reconstructed and if some errors in the design of the machine were tolerated.[79] As things stand now, however, automated vehicles could barely be perceived as ordinary parts of our lives that would be widely accepted by people.[80]

Is a Shift away from the Liability of the Human in the Driver's Seat viable?

The previous legal treatment of new technology has been investigated by Glancy et al.[81] New technologies have been addressed in a number of ways in terms of policy. Technologies that have been regarded either dangerous or coming with advantages have attracted quick policy reactions. In contrast, some innovations, including airplanes have not caused a rush to policymaking.[82] Frequently, feedback to new technology has developed gradually. That is explained by the evolution of technology and changed assumptions.[83] For instance, speed limits for cars were reviewed from time to time. Reactions by policymakers have made the most of increasing understanding of the given technology and the peril it poses – one may think of the rules targeting driving when intoxicated.[84] Such policy outcomes tend to follow changes in attitudes in terms of the advantages and disadvantages of new innovations.[85]

In terms of automated vehicles, it may, however, be controversial if in contrast to an earlier state of affairs the bearer of liability would be a third party instead of the user.[86] Current ideas of how liability is allocated between car makers and users may be difficult to get rid of. It may take time before it gets settled what are the features of vehicle behavior that make the producer liable on the one hand and on the other where the car's actions would make the user liable. While this uncertainty prevails, it might be advisable not to introduce any changes to the liability scheme.[87]

The UK inquiry noted that people might be more approving of risks that are known, up to the individual or come with understandable advantages. Novel and unusual risks that are imposed on people are less acceptable.[88] It was further noted that the public does not find deaths and injuries caused by bad human drivers acceptable. There is an intense desire to introduce criminal penalties for such conduct.[89] Research seems to indicate that a substantial number of the public thinks that the safety of automated vehicles should be comparable to air and rail travel which would arguably be reflected in risk acceptability in terms of novel transport over which people do not have the control.[90] The UK report seemed mindful of the potential public attitudes when it discussed aggravated offences for death or serious injury as a result of an ADSE's wrongdoing. With aggravated offences the underlying behavior is already criminal, but the outcome invites a harsher treatment. This could be the case if for instance dangerous driving causes death. Aggravated driving offences

79. Gless et al., 2016, pp. 430-431.
80. Ibid., p. 433; Kaiafa-Gbandi, 2020, p. 316.
81. Glancy Dorothy J. et al., 'A Look at the Legal Environment for Driverless Vehicles,' National Academies of Sciences, Engineering, and Medicine 2016.
82. Ibid., p. 15.
83. Ibid., p. 16.
84. Ibid., 2016, p. 16.
85. Ibid., p. 16.
86. Glancy et al. 2016, p. 42.
87. Ibid., p. 42.
88. For people, numeric details do not seem to be decisive, Law Commissions 2022, p. 56.
89. Ibid., p. 56.
90. It could be unfair if vulnerable groups would as a result of automated vehicles see a shift of risks to their disadvantage from the car occupants. A feeling of fairness will arguably have an impact on how the public perceives the risk involved, see Ibid., pp. 60-61.

are widely supported by the public. It was argued that when illegal behavior by an ADSE causes death or serious injury the public expects that this is displayed in the punishment.[91] Arguably, applying a regulatory offence would not be adequate.[92] The Swedish Memorandum acknowledged that it might be experienced as unfair from the citizens' perspective if acts which would entail criminal responsibility for a manual driver would remain outside the reach of sanctions.[93] In sum, it seems that a sense of unfairness may not be remedied by a regulatory approach that no longer holds the human user liable.

This conclusion is reinforced by psychological evidence which points to the retributivist nature of humans.[94] In Finland[95] there was one accident involving a Tesla model 3 car crashing into another vehicle. The autopilot had been engaged. The prosecutor accused the defendant of having failed to keep an eye on the automated driving and the traffic which then led into a crash with another vehicle.[96] After the accident, the victim wanted to know what the defendant "was thinking of" to which the defendant had said that it was "the car which did the thinking," not him.[97] The court, however took the position that it was not proven that the defendant had completely depended on the driver assistance system.[98] Nonetheless, this case may be indicative of how cases concerning SAE level 3 vehicles might be received in courts and more widely by the public. This seems to be supported also by the survey findings of Awad et al.—when both the human and machine drivers of a vehicle make mistakes while the vehicle is under their control, it is the machine which gets to shoulder less blame.[99]

On the whole, a wide-ranging impunity in terms of AI systems may not be ideal if the community finds it troublesome that there is a lack of individual accountability.[100] From a legal point of view, arguably, other areas of law may be better suited than criminal law for introducing requirements for using and developing smart machines. Yet letting other areas of law exclusively address the conundrum may not be the best possible outcome. A society's confidence in rules might be upset if technological advancements end up infringing rights laid down by law.[101]

91. It was noted that also moral reasons speak for liability, see Law Commissions 2022, p. 220.
92. Ibid., p. 222.
93. Ds 2021:28, p. 230.
94. See the discussion in Danaher John, 'Robots, law and the retribution gap,' *Ethics Inf Technol* 2016, p. 299ff.
95. Regarding the close legal and cultural ties between Finland and Sweden see Lahti Raimo 2022, p. 12ff.
96. South Savo district court, tuomiolauselma 21/115548, asianumero R20/968, 13.04,2021, pp. 1-2.
97. The court, however, bought the defence's argument that the defendant had said this in an attempt to defuse a potentially explosive situation.
98. South Savo district court, tuomiolauselma 21/115548, asianumero R20/968, 13.04.2021, p. 10.
99. Awad Edmond et al., 2020, pp. 139-140.
100. It has been argued that not holding individuals accountable or undermining individual accountability may not seem compatible with the German legal system. Assigning individual responsibility is important to the way how we perceive each other and ourselves, see Beck 2020, p. 48.
101. Also, from the point of view of the victim the infringement of his rights may not be fully made up by material redress. Therefore, having no avail to criminal law measures may not seem convincing. A kind of normative message is sent by the society via criminal law rules. In terms of smart machines that message would be absent if no criminal responsibility was imposed. Arguably, other means of sending normative messages would not be on a par with criminal law measures, Beck 2020, p. 49; '[T]he rule of law can be undermined if legal systems fail to align with these intuitive judgments', see Danaher John, 'Robots, law and the retribution gap,' *Ethics Inf Technol* 2016, p. 308; A financial penalty would not seem to be a fair penalty for a human rights infringement, see Mcallister Amanda, 'Stranger than Science Fiction: The Rise of A.I. Interrogation in the Dawn of Autonomous Robots and the Need for an Additional Protocol to the U.N. Convention Against Torture', *Minnesota Law Review* 2017, p. 2563; King, Thomas and Aggarwal, Nikita and Taddeo, Mariarosaria and Floridi, Luciano, Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions (May 22, 2018). Available at SSRN: https://ssrn.com/abstract=3183238 or http://dx.doi.org/10.2139/ssrn.3183238, last visited on August 19, 2022, p. 16; See also Kaiafa-Gbandi, 2020, pp. 323-327.

Concluding remarks

It seems that in various areas of application of AI (e.g., automated vehicles) absolving the human user of blame may seem premature. This conclusion owes to the idea that smart machines could barely be perceived as ordinary parts of our lives that have been embraced by our societies.

With a view to avoiding responsibility gaps it has been suggested that a responsible person could be designated. In terms of automated vehicles that situation could materialize where the driver needs to comply with a continuous oversight duty. The rationale for this approach could be found in the developmental phase of the new technology—the risks involved appear to be still relatively high or uncertain. While imposing the aforementioned duty on the driver could be labeled as strict liability, it may be noted that punishing unconscious negligence is nothing new in terms of traffic violations.

The use of strict liability has been justified by public protection objectives. It cannot be ruled out that it may be an appropriate objective when it comes to the application of AI in the sense that the sort of reasoning that could be found in *Empress Car* and the decision of the Finnish Supreme Court might be useful in future cases regarding AI. That conclusion is further supported by the Proposal for a Regulation on Artificial Intelligence which to an extent underlined human oversight and the long-established approach that would keep the user/driver in a position of responsibility (e.g. vehicles.).

All in all, a lack of individual accountability may be problematic if that is what a given community expects and may also undermine a society's confidence in rules if legal interests are violated with impunity. Therefore, a swift move to dilute human responsibility does not appear viable.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Jenny and Antti Wihurin FoundaRahasto, Makarna M. och Hj. Granfelts stipendiefond and Ella ja Georg Ehrnroothin Säätiö.

ORCID iD

Patrick S. Günsberg https://orcid.org/0000-0001-5402-4516