"The Knowledge of Causes and the Secret Motions of Things"

The Interdisciplinary and Doctrinal Challenges of Automated Driving Systems and Criminal Law

JEANNE GAAKEER[*]

The end of our foundation is the knowledge of causes and the secret motions of things; and the enlarging of the bounds of human empire, to the effecting of all things possible.[1]

I   Introduction

On October 15, 2001, a coach driver wanting to make a right turn stopped to give the right of way to a mother and her 5-year-old son on a bike crossing. After the mother had reached the other side of the crossing, she made a gesture to the driver. He accelerated and ran over the boy, who had fallen in the middle of the crossing. The boy died of his injuries. In court, the driver explained that the gesture made him assume that the boy had crossed safely. The Dutch lower, appellate, and Supreme Court found that his claim that he had based his understanding on the gesture was irrelevant, but this chapter asserts that the driver's hermeneutic (mis)understanding of the mini-narrative of the human gesture is quite relevant. Because Article 6 of the Dutch Road Traffic Act 1994, applicable to traffic accidents resulting in grave bodily injury or death, is based on *culpa lata*, i.e., behavior less careful than that of the average person, the presumption of innocence allows a defendant to plead not guilty based on his or her interpretation of another person's action. It appeared that the disastrous consequence of the boy's death occasioned application of a stricter standard, that of *culpa levis*, i.e., whether the defendant behaved

---

[*] I am grateful to Sabine Gless for including me in this exciting project.
[1] Francis Bacon, *The New Atlantis* (USA: Project Gutenberg, 2008), www.gutenberg.org/files/2434/2434-h/2434-h.htm.

335

as the most careful person possible.[2] As Ferry de Jong suggests, when it comes to determining *culpa*, guilt, and *dolus*, intentionality, in any specific criminal case, a "hermeneutics of the situation"[3] is required to gauge whether or not *actus reus* and *mens rea* can be established. In this chapter, hermeneutics refers not only to the individual interpretations of actions or meaning, but also includes the criteria or framework used to produce such interpretations. A hermeneutics of the situation stresses the connection between this process of meaning-giving and the situation in which the process occurs.[4] This process is difficult enough in traffic accidents involving traditional cars, and it will become even more difficult if the car is a robot.

An autonomous vehicle is a robot, and a robot is understood here as "an engineered machine that senses, thinks, and acts."[5] In view of the increased use of automated vehicles, referred to in the chapter as Automated Driving Systems (ADS), the need for a hermeneutics of the situation has become even more acute. When ascertaining the degree of criminal fault when ADS are involved in traffic accidents, we have to face the unpleasant truths that so far legislation lags behind and current versions of legal codes may fall short. Criminal law concepts dealing with intent and causality therefore need a new, careful scrutiny, because ADS have their own hermeneutics, one which is not easily comprehensible to the driver. ADS hermeneutics are based on their programming, i.e., their algorithms, and this introduces novel understandings of what it means to act – hermeneutical as well as narratological.[6]

In addition to drivers of ADS, legislators may also find the logic of new technologies fuzzy. The question of hermeneutically understanding technology at the legislative level is outside the scope of this chapter, and limited space does not permit me to elaborate. It can be noted that any

---

[2] Dutch Supreme Court, Decision of January 17, 2006, ECLI:NL:HR:2006:AU3447. European judicial decisions that have a European Case Law Identifier (ECLI) can be accessed via the European e-justice portal, see European Union, "European e-justice," https://e-justice .europa.eu.

[3] Ferry de Jong, "The End of Doctrine? On the Symbolic Function of Doctrine in Substantive Criminal Law" (2011) 7:3 *Utrecht Law Review* 8 at 44, n. 141, referencing Antoine Mooij, *Intentionality, Desire and Responsibility: A Study in Phenomenology, Psychoanalysis and Law* (Leiden, Netherlands: Brill, 2010) at 39–45 ["End of Doctrine"].

[4] See Antoine Mooij, "*Psychiatry as a Human Science: Phenomenological, Hermeneutical and Lacanian Perspectives* (Leiden, Netherlands: Brill, 2012) at 156.

[5] For this understanding of robot, see Chapter 6 in this volume.

[6] Narratology refers to the theory and study of narrative as story and storytelling, while the latter is the narrative representation of human actions, events, and happenings.

legislative choice regarding ADS in criminal law will influence future criminal charges, which are themselves always already mini-narratives of forms of reprehensible human behaviour, *mala prohibita*.[7] Both future legislation and pending concrete cases are in need of an informed hermeneutics of the situation, disciplinary and factual, not least because hermeneutic misunderstanding may be an impediment to the right to a fair trial.

Many disciplines were already involved in the development and construction of ADS before jurists became involved. The difficulties of how to interpret and understand the disciplinary other may easily lead to miscommunication when artificial intelligence (AI) experts who are not jurists must deal with jurists who are not AI experts.[8] In addition to problems of translation between disciplines, responsibility gaps may occur, "circumstances in which a serious accident happens and nobody can be reasonably held responsible or accountable due to the unpredictability or opaqueness of the process leading to the accident," technological opaqueness included.[9] For example, in 2020, a former member of the EU Parliament, Marietje Schaake, had a conversation with an entrepreneur. The entrepreneur told her that one of his engineers working on the design of ADS had asked him who he would prefer to be killed in case of a collision involving an ADS, either a baby or an elderly person, because such options had to be built into the software.[10] This brings to mind the ethical-philosophical thought experiment called the "Weichenstellersfall" or trolley problem. A train runs out of control and will kill hundreds of people in a nearby train station unless it is diverted to a side track, but on that track there are five workmen who will be killed as a consequence. What should be done? Do you divert the train or not? Even more complicated is the problem's elaboration in the fat man example; what if you are on a bridge and the only way to stop the train is to kill a fat man next to you and

---

[7] For the idea of the criminal charge as a mini-narrative, see Jeanne Gaakeer, "The Criminal Charge: A Narratological Bow Tie?" in Monika Fludernik & Frank Schäfer (eds.), *Erzählen und Recht* (Narrative and Law), 12 Faktuales und Fiktionales Erzählen (Baden-Baden, Germany: Ergon, 2022) 129.

[8] AI is understood here to include "neural networks engaged in deep learning"; see Chapter 7 in this volume.

[9] Filippo Santoni de Sio, "Ethics and Self-Driving Cars: A White Paper on Responsible Innovation in Automated Driving Systems" (Delft University of Technology, 2016) ["White Paper"] commissioned by Rijkswaterstaat for the "knowledge agenda automated driving," at 20. The term "responsibility gap" was coined by Andreas Matthias in Andreas Matthias, "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata" (2004) 6:3 *Ethics and Information Technology* 175.

[10] Economy Section, "NRC Handelsblad" (February 7, 2020) at 8.

push his body on to the track to stop the train?[11] Translated to the topic of ADS, when there is imminent danger, the human driver and/or the ADS have to decide between two evils and choose to kill either one person or the other(s). Any human driver killing one individual in order to save the other(s) will be acting unlawfully, but would that also be acting culpably? Furthermore, if a democratic state under the rule of law can never weigh the life of one citizen against the other and prohibits any distinction on the basis of age, gender, and sex, why would we allow an engineer to do just that when programming an ADS? Understanding our fellow human beings and their actions is difficult enough, but understanding, let alone arguing with, an algorithm not of one's own design is even more so. Technological advances in driving may be intended to reduce the complexity of the human task of driving a vehicle in contemporary traffic – the technological narrative of progress – but may in fact complicate it if such innovation demands that the human be on the alert for any surprise in the form of an error in the algorithmic and/or computational system, causing the vehicle to deviate from its intended course. While research is being done on how human drivers understand and use specific types of ADS, the current human driver-passenger may be hermeneutically challenged. How and when does she recognize that she needs to resume control?

While criminal law does not solely represent the pursuit of moral aims, new AI technologies force us to consider ethical issues in relation to hermeneutical and narratological ones, and to grapple with the criminal liability of ADS. To this end, the chapter incorporates different interdisciplinary lenses, including narratology. The chapter is inspired by the epistemological claim on human knowledge and progress voiced in Francis Bacon's utopian narrative *The New Atlantis*, because the fundamental philosophical questions "What is it? What do you mean? How do you know?" apply in technological surroundings as much as in criminal law surroundings. The actors involved have to be able to clearly express their stories, paying careful attention not only to *what* they are saying and claiming, but also to *how* they tell their stories.[12] These ontological, hermeneutical, and methodological questions are therefore narratological questions as well.

---

[11] Hans Welzel, "Zum Notstandsproblem" (1951) 63:1 *Zeitschrift für die gesamte Strafrechtswissenschaft* 47; Judith Jarvis Thomson, "The Trolley Problem" (1985) 94:6 *Yale Law Journal* 1395.

[12] On the narratological distinction between narrative as story and as discourse, see Gerald Prince, *A Dictionary of Narratology*, rev. ed. (Lincoln, NE and London, UK: University of Nebraska Press, 2003), Discourse ("the expression plane of narrative as opposed to its content plan or story").

In Section II, this chapter addresses the interdisciplinary issues of integrating knowledge, translating between disciplines, and responsibility gaps, as a prolegomena for Section III, which focuses on criminal liability. In Section IV, the human–robot/ADS interaction is discussed, in the context of issues raised by the concept of *dolus eventualis*. To conclude, Section V returns to the need for a hermeneutics of the situation that adequately addresses ADS challenges.

II   Interdisciplinary Observations on the Interrelation of Technology and Law

II.A   Whose Department?

The legal implementation of technology is too important to leave to technologists alone. This chapter therefore turns to philosophical thought on technology, in part to prevent us from falling into the trap of Francis Bacon's *idola tribus*, i.e., our tendency to readily believe what we prefer to be true.[13] The *idola tribus* makes us see what our rationalizations allow. This approach is the easy way out when we do not yet fully understand the effects and consequences of new technologies, but the moment is not far away when ADS becomes fully capable of independent, unsupervised learning, and we should consider Samuel Butler's visionary point on the side-effect of machine-consciousness, i.e., "the extraordinary rapidity with which they are becoming something very different to what they are at present."[14] When that happens, who or what will be in control?

An epistemology based on algorithmic knowledge, while helpful in many applications to daily life, runs the risk of introducing forms of instrumentalism and reductionism. Behind such "substitutive automation" is the "neoliberal ideology … [in which] dominant evaluative modes are quantitative, algorithmic, and instrumentalist, focused on financialized rubrics of productivity."[15] The greater the complexity of the issue, the greater the risks posed by algorithmic knowledge. Scientific dealings in these modes of analysis often disregard the fact

---

[13] Joseph Devey (ed.), *The Physical and Metaphysical Works of Lord Bacon, Including The Advancement of Learning and Novum Organum* (London, UK: George Bell & Sons, 1901) at 209.

[14] Samuel Butler, *Erewhon* (Harmondsworth, UK: Penguin, 1954) at 164.

[15] Frank Pasquale, "Professional Judgment in an Era of Artificial Intelligence and Machine Learning" (2019) 46:1 *Boundary* 2 ["Professional Judgment"] at 1 and 2.

that a human being is the source of the data, both as the object of the algorithms used in technologies when data is gathered to run the device, and as the engineer and designer who decides what goes into the programming process. Human fallibility is often disregarded, but ontological perfection either of humans or technologies is not in and of this world. While both human and AI learn by iteration, their individual awareness of past and present danger is not identical, or should we say, identically programmed.

Some Dutch examples may illustrate the difficulties in relying exclusively on algorithmic knowledge. In 2018, the advanced braking system of a Volvo truck failed because the camera system did not recognize a stationary truck in front of it in the same lane.[16] In the subsequent crash into the back of another truck, the driver of the Volvo was crushed to death. In a 2017 case, the warning system of a 2014 model Tesla failed to respond to another vehicle that changed lanes, the Tesla did not reduce its speed in due time, and it hit the side of the other vehicle. The manufacturer admitted that the 2014 model worked well when it came to detecting vehicles right in front of the Tesla, but not when these vehicles made sudden moves.[17] But that is not an uncommon event in traffic, is it?

The examples show that data-driven machines run the risk of incorporating forms of "epistemological tyranny."[18] The human is reduced to the sum of its "dividual" parts, selectively used depending on its user's needs.[19] Our making sense of the relations between individuals and their machines is then reduced to connecting the dots. If manufacturers focus on the development of new technologies rather than on the legal frameworks within which their products are going to be handled, any opacity as far as product information is concerned can lead to someone, somewhere, avoiding compliance with the law. We should therefore probe the "*narrative* of computationalist supremacy."[20] The humanities can help provide guidance at the meta-level of juridical-technological

---

[16] Netherlands, Dutch Safety Board, *Wie Stuurt? Verkeersveiligheid en automatisering in het wegverkeer* (Who's Driving? Traffic Safety and Automation) (The Hague, Netherlands: Dutch Safety Board, 2019) at 24.

[17] Ibid. at 31–45.

[18] "Professional Judgment", note 15 above, at 15.

[19] Gilles Deleuze, "Postscript on Control Societies" in Thomas Levin, Ursula Frohne, & Peter Weibel (eds.), *CTRL [SPACE] Rhetorics of Surveillance from Bentham to Big Brother*, 1st ed. (Cambridge, MA: The MIT Press, 2002) 317 at 319.

[20] "Professional Judgment", note 15 above, at 30, n. 2 (emphasis in the original).

discourse, because behind any form of "algorithmic imperialism,"[21] there is also linguistic imperialism that prioritizes one language of expertise above the other.[22]

Under the influence of Enlightenment thought, the stereotypical or stock story of modern technology, its constitutive narrative, founded as it is in the natural sciences, has been the narrative of human progress.[23] Its darker side-effects have often been pushed into the background until something went seriously wrong. But it is a mistake to regard technology "as something neutral."[24] If we look upon technology as production only, we may be reduced to Deleuzian dividuals, ready to be ordered by others, be they machines or humans, both in technology and law; then "'[t]he will to mastery' will prevail and we have to wait and see who gets in control at the level of production."[25] While the heyday of legal positivism is behind us, its referential paradigm may well resurface, if for lack of information or understanding we all too readily accept at face value what is held before us as technology. The consequence may be uninformed and unethical applications of technology, without proper legal protection of the humans impacted by it.

This chapter does not promote Luddism. It does, however, highlight the risks involved in a positivist view of both law and technology, i.e., the value-free, unmediated application of any form of code, as opposed to the value-laden human enterprises that they are. As Lawrence Lessig put it, "Code is never found; it is only ever made, and only ever made by us."[26] Technology should not be put to use for the simple reason that it is available, and one risk of modern technologies is that if it can be done, somewhere, someone, at some point in time, will actually do it, whatever the consequences. This attitude is brilliantly and cynically

---

[21] Ibid. at 16.

[22] See James Boyd White, *Living Speech: Resisting the Empire of Force* (Princeton, NJ: Princeton University Press, 2006).

[23] For the idea of constitutive narratives in relation to law, see Robert Cover, "Nomos and Narrative" (1983) 97:1 *Harvard Law Review* 4 ["Nomos and Narrative"] at 4–68.

[24] Martin Heidegger, "The Question Concerning Technology" in Martin Heidegger, *The Question Concerning Technology and Other Essays*, translated by W. Lovitt (New York, NY: Harper & Row, 1977) 3 at 4. It should be noted that Heidegger's career was severely tainted by his association with the National Socialists during his rectorate of the University of Freiburg. Despite this controversial aspect, he is widely regarded as one of the greatest philosophers of hermeneutics.

[25] Ibid. at 5.

[26] Lawrence Lessig, *Code: And Other Laws of Cyberspace, Version 2.0* (New York, NY: Basic Books, 2006) at 6.

voiced in Tom Lehrer's 1965 song "Wernher von Braun": "'Once the rockets go up, who cares where they come down? That's not my department,' says Wernher von Braun."[27] Careful attention regarding the what, the how, and the why of ADS technology is required. The what of the algorithm, the logic of the if … then, does not coincide with the how of its juridical-technical implementation, let alone the how of its technical discourse. This is no small matter if we think of the if … then structure of the criminal charge in terms of punitive consequences for human behavior involving ADS, and the narratives a defendant would need to steer clear of criminal responsibility.

II.B   The Need to Integrate Knowledge

Mono-disciplinary approaches reinforce scientific dichotomies that preclude the necessary risk assessments. They bring us back to the *Erklären-Verstehen* controversy, as it is called in the nineteenth-century German philosophical tradition, to the concept of restricting explanations to the natural sciences, because explanation (*Erklären*) could only pertain to facts, whereas the humanities could only attribute meaning or hermeneutic understanding (*Verstehen*). This dichotomy has had far-reaching implications for the epistemological differentiation of knowledge into separate academic disciplines, with each discipline developing its own language and methodology, outlook, goals, and concepts, and each discipline functioning in a different cultural and social context of knowledge production. The interdisciplinary approach advocated here can show that in all epistemological environments, "[d]isciplinary lenses inevitably inform perception."[28] An interdisciplinary approach also calls for an appreciation of the fact that any discipline's or field of expertise's narratives cannot be understood other than within their cultural and normative universe, the *nomos* of their origin and existence.[29]

To see the connection between ADS technology and narratology, we could ask what the new technologies' rhetoric, scripts, and stock stories have been so far, and specifically, what the main narrative thrust of technology is and what it means for the non-specialist addressee. Any field of knowledge "must always be on its guard lest it mistake its own linguistic

---

[27] Tom Lehrer, "That Was the Year That Was" (1965).
[28] Siri Hustvedt, *The Shaking Woman or the History of My Nerves* (London, UK: Picador, 2010) at 28.
[29] "Nomos and Narrative", note 23 above, at 4–68.

conventions for objective laws."[30] Debate is essential, and engineers and jurists alike need guidance regarding the production and reception of narratives in their respective fields. One such form of guidance is Benjamin Cardozo's claim that legal professionals need to develop a linguistic antenna sensitive to peculiarities beyond the level of the signifier, because the form and content, the how and the what of a text, are interconnected.[31] Concepts from narratology can assist to accomplish this task. All professionals benefit if they learn to differentiate between, first, narrative in the sense of story or *what* is told, and discourse of *how* it is told. For jurists working in criminal law, it is important, second, to realize that story comprises both events, understood here as either actions or happenings, and the characters that act themselves or get involved in happenings, and that all of this occurs in specific settings that influence meaning.

Precisely because disciplinary lenses influence us, translating between collaborating disciplines must be undertaken. To the legal theorist James Boyd White, interdisciplinarity is itself a form of translation. He claims that resolving the tensions between disciplines "always involves the establishment of a relation between two systems of language and of life, two discourses, each with its own distinctive purposes and methods, its own ways of constructing the social relations through which it works, and its own set of claims, silences, and meanings."[32] At the core of translation as a mode of thought, then, is the claim that we should be alert to the possibilities and limitations of any professional discourse. This point illuminates the possibilities and limitations of any disciplinary language of expertise, limitations tied to the context of claims of meaning, and to the cultural and social effects of specific language uses. Translation requires that we address the fundamental difference between the narrative and the analytical, between "the mind that tells a story, and the mind that gives reason" because "one finds its meaning in representations of events as they occur in time, in imagined experience; the other, in systematic or theoretical explanations, in the exposition of conceptual order or structure."[33] When

---

[30] Italo Calvino, "Two Interviews on Science and Literature" in *Italo Calvino, The Uses of Literature*, translated by P. Creagh (New York, NY: Harcourt Brace Jovanovich, 1987) at 45.

[31] See Benjamin Cardozo, "Law and Literature" (1925) 14 *Yale Review* 699.

[32] James Boyd White, "Establishing Relations between Law and Other Forms of Thought and Language" (2008) 1:3 *Erasmus Law Review* 1 at 9, www.elevenjournals.com/tijdschrift/ELR/2008/3/ELR_2210-2671_2008_001_003_002.pdf.

[33] James Boyd White, *The Legal Imagination: Studies in the Nature of Legal Thought and Expression* (Boston, MA: Little, Brown & Co., 1973) at 859.

transposed to the subject of conceptual thought, the need for attention to language and narrative becomes acute. What, to start with, is "a concept"? White found "concept" a problematic term, because the underlying premise is once again the referentiality of language, one that implies transparency of the semantic load of a concept in one disciplinary language and, following this, unproblematic translation of a concept into another. Such a view is imperialistic, based as it is on the supposition that the "conceptual world … is supposed to exist on a plane above and beyond language, which disappears when its task is done."[34]

One central example of translation in the context of human–ADS interactions is the concept of driver, currently presumed to be a human driver. In a present with current levels of ADS development, and in a future of full ADS automation, a legal concept of the driver based on a human is no longer appropriate. Feddes suggests that "the human is a passenger, the automation is the legal driver."[35] If this is correct, attribution of legal responsibility in human–ADS interactions would require ADS to be able to handle any situation that crops up.

A Dutch case on the concept of driver illustrates arguments regarding who the driver is in a human–ADS interaction. The driver of a 2017 Tesla Model X was fined €230 in an administrative sanction for using his mobile phone hands-on while driving.[36] Before the county court, he claimed that because the autopilot was activated, he could no longer be legally considered the driver, and therefore the acts of driving and using a hands-on phone did not constitute the simultaneous act prohibited in Article 61A of the Rules on Traffic Regulations and Traffic Signs 1990.[37] This narrative did not save the day. The county court found the defendant's appeal unfounded because Article 1 of the Road Traffic Act 1994 applied. The defendant had stated that while seated on the driver's chair with the autopilot activated, he regularly held the steering wheel, but he did this because the system disengages itself if the driver does not react

---

[34] James Boyd White, *Justice as Translation: An Essay in Cultural and Legal Criticism* (Chicago, IL: University of Chicago Press, 1990) at 31 and 36.

[35] Gerben Feddes, "Towards the Legal Admission of Connected Automated Vehicles," Paper EU-TP1330 delivered at the 25th ITS World Congress Copenhagen (September 17–21, 2018) 1 at 5.

[36] County Court Midden-Nederland, Decision of November 22, 2018, ECLI:NL:RBMNE: 2018:5707 [ECLI:NL:RBMNE:2018:5707].

[37] Per the Rules on Traffic Regulations and Traffic Signs 1990 [Rules on Traffic], Art. 61A, the legal driver is: "A person driving a motor vehicle, moped, motor assisted bicycle or disabled person's vehicle equipped with an engine may not hold a mobile phone while driving."

after the three auditory warnings from the vehicle when it notices that the driver is not holding the wheel.[38] He was found to be the legal driver of the vehicle and not a passenger, in part because drivers are "all road users excepting pedestrians" according to Dutch law.[39] Like the Netherlands, many legal systems lack a codified definition of the term "driver," which leads courts to define the term in context.

The defendant's other argument in this case, that Dutch legislation should be amended to provide a definition, did not help the defendant either, because in criminal cases future-oriented contextual interpretation is prohibited. On appeal, the defendant introduced a new element to his narrative, that a driver using an autopilot is similar to and should be treated like a driving instructor. Since a driving instructor is not the actual driver, he or she is allowed to use a mobile phone hands-on. This narrative forced the Court of Appeal to elaborate on the doctrinal distinction made in the Road Traffic Act 1994 and the Traffic Rules and Signs Regulations 1990 between the actual driver and the legal driver. Article 61A of the Traffic Rules and Signs Regulations 1990, the regulations used for the administrative charge against the defendant, pertained to the actual driver, not to the instructor or examiner. Activating and using the autopilot, as the defendant had done, made the defendant the actual driver, as his vehicle was not a fully automated ADS. Per this reasoning, Article 61A applied. The Court of Appeal upheld the judgment.[40] Under this reasoning, there is nothing automatic in autopilots yet!

A final, comparative question regarding translation is whether the process of ADS construction reflects unconscious biases. Suppose an ADS is of US American design. Surely the designer had US American law at the back of his mind during construction? Does such a vehicle fully comply with the demands of civil-law European systems and the mindsets of European users? An interdisciplinary approach regarding technology and law compels us to think through incompatibilities, while at the same time urges us to integrate their disciplinary discourses as much as possible. Rather than continuing a "'black box' mentality,"[41] we should promote

---

[38] ECLI:NL:RBMNE:2018:5707, note 36 above.

[39] Rules on Traffic, note 37 above, s. 1.

[40] Dutch Court of Appeal Arnhem-Leeuwarden, Decision of July 31, 2019, ECLI:NL:GHARL: 2019:6122.

[41] Luciano Floridi, Josh Cowls, Monica Beltrametti *et al.*, "AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations" (2018) 28:4 *Minds & Machines* 689 at 692, "a 'black box' mentality, according to which AI systems for decision-making are seen as being beyond human understanding, and hence control."

"technologies of humility,"[42] to preclude technological languages from imposing their conceptual framework to the exclusion of other languages.

II.C   Mind the Gap

As noted above, a responsibility gap arises when a serious accident happens but nobody can reasonably be held responsible. Responsibility gaps can arise because of the gaps between disciplinary fields. An example of minding the disciplinary gaps is Santoni de Sio's attention to ethical issues, in which he urges integration of different disciplines. He observed that the Dutch Ministry of Infrastructure and Environment divides ethical issues in ADS into three levels: the operational level concerning the programming of automated vehicles; the tactical level of road traffic regulations; and the strategic aspect of how to deal with the societal impact of ADS.[43] For ADS, integration "should be done in such a way that 'meaningful human control' over the behaviour of the system is always preserved."[44] The simple fact that a human is present is not in itself "a sufficient condition for being in control of an activity."[45] This is the case because of the complexity of all the causal relations and correlations involved, and because "meaningful" control is not equivalent to "direct" control, i.e., when the driver directly controls the ADS's full operation. Confusing meaningful and direct control can easily lead to either over-delegation, as when the driver of an ADS overestimates the vehicle, or under-delegation, where the driver overestimates his or her own driving capacities in an ADS context.[46] The need to clearly define the scope of the driver's actual freedom to act is also inextricably connected to the notion of volition in criminal law.

III   Criminal Liability

III.A   Freedom to Act?

Human autonomous agency is inextricably connected to consciousness and to the capacity for rational thought. With these come free will, manifesting in criminal law, first as the self-determination to deliberately do

---

[42] Ibo van de Poel, "An Ethical Framework for Evaluating Experimental Technology" (2016) 22:3 *Science & Engineering Ethics* 667 at 668, referencing Sheila Jasanoff, "Technologies of Humility: Citizen Participation in Governing Science" (2003) 41:3 *Minerva* 223.
[43] "White Paper", note 9 above, at 5.
[44] Ibid. at 8.
[45] Ibid. at 11.
[46] Ibid. at 14–15.

the right thing and abstain from what is wrong, e.g., *mala per se* such as murder, and *mala prohibita* or what the law prohibits, and second as the criterion for assigning legal personhood. When it comes to attributing criminal liability, the first requirement is *actus reus*, the voluntary act or omission to act that the law defines as prohibited. Historically, the free will necessary for a voluntary act has been defined in numerous ways. It can mean that man is free to decide to go either left or right, even if there is no specific reason to do either. One has freedom to act if one is able to do whatever one decides, the *liberum arbitrium indifferentiae*.[47] Free will can also be seen when one is free to decide not to act at all. This is the precursor and precondition of the legal freedom to act in that it presupposes the mental ability to decide whether or not to do this, that, or the other.[48] The fact that man is aware of the fact that he has a will is not deemed enough, because being conscious of something is not evidence of its existence.

What are the necessary and sufficient conditions of a voluntary act in the context of ADS, and what are the legal consequences of those conditions? The lack of free will is still widely regarded as the axe at the root of the criminal law tree. The question today in human–robot relations is whether or not free will and forms of technological determinism can be reconciled, theoretically and practically. Is free will compatible with empirically provable determinants of action? If so, then free will is perhaps compatible with machine-determined action, and therefore legal causality. The necessary condition for free will is that an actor, in doing what he did, could have decided otherwise. In the law, we normally start from the premise that free will is a postulate that goes for the majority of ordinary human beings opposed to an empirically provable fact, because statistically speaking that is usually the situation. This approach leads to the traditional position that those suffering from mental illness are not free, and hence not or only partly responsible. The law's beginning assumption of free will also leads to the impossibility of punishing those about whom one cannot say anything other than we do not know whether their will was hampered or not. Practically speaking, free will is established when a state of exception, e.g., insanity in humans, does not occur.

---

[47] Pierre Bayle, *Dictionnaire historique et philosophique*, vol. II (Amsterdam, Netherlands: Compagnie des Libraires, 1734) at 466.

[48] Julien Benda & Raymond Naves (eds.), *Voltaire Dictionnaire Philosophique* (Paris, France: Garnier, 1961) at 277, "Vous êtes libre de faire, quand vous avez le pouvoir de faire."

Two opposing views regarding the application of these ideas to ADS could be entertained. One is that if an ADS is an agent capable of learning in the sense of adapting its actions to new information, an ADS could be held criminally responsible, with or without attributing consciousness of the human type, because the algorithmic reasoning skills and autonomy of the ADS would suffice. Second, if charges are brought against the human driver, one could argue that an ADS provides a defense based on the state of exception approach to free will discussed above. The human driver does not know the mind of the ADS and cannot probe the technological sanity of an ADS, partly because the ADS is a device programmed to act in response to its environment, but not by the driver.

Both views are connected to the question of a possible form of legal personhood for AI, another condition for the imposition of legal responsibility. As a status conferred by law on humans and entities such as corporations, legal personhood is a construct. In everyday life, it is relatively easy to recognize a fellow human being if you meet one. We then recognize the rights and responsibilities of that independent unit, and we distinguish among different entities with legal personhood, e.g., between a toddler without and an adult with legal obligations. Things are already more difficult regarding artificial persons such as corporations, in terms of the information required to assess what the artificial person's rights and obligations are, and the inquiry becomes more fraught regarding ADS.[49] Another issue is that as a matter of legal doctrine, most countries have a closed system of legal personhood. Adding to it may not be as easy as, e.g., the European Parliament thought, when in 2017 it spoke about personhood in the form of an "electronic personality" for robots[50] without explaining which form it could or should take. The European Commission then declined granting such legal status to AI devices.[51]

---

[49] For a comparison of criminal responsibility for corporate entities and robots, see Chapter 4 in this volume.

[50] European Union, European Parliament, Civil Law Rules on Robotics: European Parliament Resolution of 16 February 2017 with Recommendations to the Commission on Civil Law Rules on Robotics, P8_TA(2017)0051 (EU: Official Journal of the European Union, 2018), para. 59(f).

[51] European Union, European Commission, Artificial Intelligence for Europe, COM(2018) 237 final (Brussels: European Commission), PE 621.926, s. 2.1.22.1.1, https://ec.europa.eu/transparency/documents-register/api/files/COM(2018)237_0/de00000000142394?rendition=false; see also European Union, European Commission, Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions – Artificial Intelligence

The issues of legal personhood and voluntariness are related. Voluntariness of the *actus reus* of any criminal charge is an issue for ADS. We assume that humans have volition because they do most of the time, and so the law does not always explicitly address the question of human volition. However, voluntary participation in an action is intimately connected to the Enlightenment model of thought that has individual autonomy at its heart and informs our current understandings of law. The requirement for voluntariness therefore prompts the issue of legal personhood to return with a vengeance, because the *actus reus* of a criminal charge, as the outwardly visible activity subject to our human understanding and judgment, is understood to be one committed by a legally capable and responsible person, unless otherwise proved. In short, the basic proposition of criminal law is that if one has legal personhood, one can be held responsible, if there is sufficient evidence and if the *actus reus* is accompanied by *mens rea*, the guilty mind. Legal personhood and voluntariness are elements that therefore remain inevitably entangled in any discussion of criminal liability and ADS.

III.B   *Which Guilt and Whose Guilty Mind?*

Mens rea, the requisite mental state that accompanies the *actus reus*, is required for criminal responsibility, and a precise articulation of *mens rea* is in turn required by substantive due process. But because criminal law regarding ADS is currently under-developed, we should be even more aware than usual of the doctrinal differences regarding *mens rea* terminology at different levels. In particular, when comparing legal systems, legal concepts applicable in common law settings cannot immediately be translated to civil law surroundings. In any discussion of *mens rea* and ADS, we are always dealing with contested definitions and fundamental differences involving the mental pictures that jurists have of their own civil law and common law concepts. Comparative research on ADS is needed, but seemingly similar concepts may be false friends.

Regarding culpability, the US American Model Penal Code[52] distinguishes between acting purposely, knowingly, recklessly, and

---

for Europe, COM(2018) 237 final (Brussels: Official Journal of the European Union, 2018); European Parliament, "Artificial Intelligence and Civil Liability," www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf.

[52] American Law Institute, Model Penal Code: Official Draft and Explanatory Notes. Complete Text of the Model Penal Code as Adopted at the 1962 Meeting of the American Law Institute at Washington DC, May 24, 1962 (Philadelphia, PA: The Institute, 1985), and subsequent revisions.

negligently, with negligence occurring when one fails to exercise the care that the average prudent person would exercise under the same conditions. Culpable criminal negligence in this framework is recklessness or carelessness that results in death or injury of another person, and it implies that the perpetrator had a thoughtless disregard of the consequences or an indifference to other people's safety. The inclusion of negligence in the Model Penal Code was controversial, because purpose, knowledge, and recklessness entail the conscious disregard of the risk of harm, i.e., subjective liability, whereas negligence does not, because the risk of harm is one that the actor ought to have been aware of, but was in fact not. Culpability as negligence is therefore often thought to result in objective, i.e., strict, liability. For many jurists, negligent criminal culpability sits uneasily with the requirement of "some mental posture toward the harm."[53] In the criminal law of England and Wales, "there is to be held a presumption … that some element of 'mens rea' will be required for conviction of any offense, unless it is excluded by clear statutory wording."[54] Various forms of *mens rea* found in statutory definitions and case law presume either: intention, direct or oblique, i.e., acting in the knowledge that a specific result will or is almost certain to occur; recklessness, either subjective, i.e., foreseen by the actor, or objective, i.e., the reasonable person threshold; or negligence, a deviation from the reasonable care standard of behavior. While recklessness resembles negligence, negligence does not coincide with recklessness.

In German criminal law, recklessness is not a separate concept. It finds a place within the concept of intention as the condition for criminal liability. Intention and negligence are the defining concepts. In this system, a negligence form of liability regarding ADS could be *dolus eventualis*, a concept which resembles the related common law concepts of recklessness and negligence, but which includes the belief that the harmful result would not occur. *Dolus eventualis*[55]

> affirms intention in cases in which the actor foresaw a possible but not inevitable result of her actions (the element of knowledge) and also approved of, or reconciled herself to, the possible occurrence of that result

---

[53] Kyron Huigens, "Virtue and Criminal Negligence" (1998) 1:2 *Buffalo Criminal Law Review* 431 at 431–432.

[54] Celia Wells & Oliver Quick (eds.), *Lacey, Wells and Quick Reconstructing Criminal Law: Text and Materials*, 4th ed. (Cambridge, UK: Cambridge University Press, 2010) at 107–108.

[55] Greg Taylor, "The Intention Debate in German Criminal Law" (2004) 17:3 *Ratio Juris* 346 ["Intention Debate"] at 348.

(the volitional or dispositional element). This is contrasted with cases in which the volitional element said to be essential to all forms of intention is missing because the actor earnestly relied on the non-occurrence of the result foreseen as possible.

Two examples may illustrate the difference between intention and negligence, and the role of *dolus eventualis*. An example of a missing volitional element was presented in a Dutch case of allegedly reckless driving. The defendant driver was driving at double the maximum speed, and the case involved a collision that killed the five passengers of the other car. The driver was charged with homicide. The Dutch Supreme Court judged him to be extremely negligent, but held that his act was not intentional as he had not consciously accepted the possible outcome of himself being killed by his own speeding, i.e., he relied on precisely the non-occurrence of an accident.[56] In a comparable German case, two persons were involved in an illegal street race which ended in an accident that killed the driver of another car who relied on the green light. The defendants were charged with murder, and the judicial debate focused on whether they had accepted the possible danger to themselves knowingly and willingly, and had been indifferent, "*gleichgültig*" as the *Bundesgericht* later called it, to the possible fate of others in case of an accident. The Berlin *Landesgericht* pronounced a life sentence, then the *Bundesgerichthof* revised the sentence on a technical matter, the *Landesgericht* then stuck to its earlier decision, and in the second revision the *Bundesgerichthof* confirmed the sentence.[57] The driver was convicted.

The dispositional element of *dolus eventualis* as indifference to what the law demands of us was developed by Karl Engisch in the 1930s, and it became the criterion to distinguish between intention and negligence.[58] In the 1980s, Wolfgang Frisch developed a risk-recognition theory. He thought of intention in terms of "an actor's realisation, at the time of acting, that a risk exists that the offence might occur, which risk the legal order regards as unacceptable."[59] Intentional action requires that the actor was aware of and deliberately created a public

---

[56] Dutch Supreme Court, October 15, 1996, ECLI:NL:HR:1996:ZD0139.

[57] See Urteil von 27.02.2017-(535 Ks) 251 Js 52/16 (8/16), Landesgericht Berlin; Bundesgerichtshof, March 1, 2018, ECLI:DE:BGH:2017:010317U4SR399.17.0; and Bundesgerichtshof, June 18, 2020, ECLI:DE:BGH:2020:180620U4STR482.19.0.

[58] "Intention Debate", note 55 above, at 355.

[59] Ibid. at 366.

wrong. Greg Taylor elaborated on Frisch's theory by means of an example in which a car driver overtaking another car on a blind corner either relies on the non-occurrence of an accident or is indifferent to the outcome. Taylor asserted that "[c]learly, by overtaking when it is not safe to do so, she creates a risk, and one which is legally unacceptable as well … Rather, the legal system condemns her conduct as unacceptable because, and as soon as, it creates a situation of danger beyond the ordinary risks of the road; it does not wait to see whether anyone is actually killed as a result of it."[60]

What issues are raised if *dolus eventualis* is applied to human driver or ADS defendants? If the foreseeability of an *abstract* risk is what is legally unacceptable, the distinction between negligence and *dolus eventualis* blurs and there is a shift in the direction of strict liability for the human driver of an ordinary car as well as for the human driver of an ADS, or the ADS itself if we accept the consequences of its self-learning. In terms of evidence, it then becomes more difficult to distinguish between intention and the advertent negligence of the driver in the Dutch example above, on the one hand, versus *dolus eventualis*, on the other. The question will then be whether we make the doctrinal move from *culpa* to *dolus eventualis* and/or strict liability in accidents involving ADS.

IV   AI and the Human: Whose Liability, Which Gap?

Societal views often differ strongly from legal decisions on the concepts of recklessness and negligence, precisely because the death of innocent people is involved. But when is an occurrence a deliberate act warranting characterization as intentional, and when is it merely an event that does not warrant criminal liability? The answer depends on the hermeneutic judicial act of evaluating facts and circumstances, and this major challenge arises in all ADS cases, not only because the information in the file may be sparse.

Identifying the *actus reus* and *mens rea* for purposes of determining wrongfulness and culpability in individual ADS cases also creates major challenges for legislators pondering policy. As Abbott and Sarch suggest, "punishing AI could send the message that AI is itself an actor on par with a human being," and "convicting AI of crimes requiring a

---

[60] Ibid. at 369–370.

mens rea like intent, knowledge, or recklessness would violate the *principle of legality*."[61] The authors develop answers to what they call the "Eligibility Challenge," i.e., what entities connected to ADS, including AI, are eligible for liability.[62] The simplest solution would be the doctrine of *respondeat superior*,[63] i.e., the human developers are responsible[64] if and when they foresee the risk that an AI will cause the death of a person, because that would be reckless homicide. The second solution is strict, no-fault liability of a defendant, and the third solution is to develop a framework for defining new *mens rea* terms for AI, which "could require an investigation of AI behavior at the programming level."[65] In court, judges could then be asked to further develop the relevant *mens rea*. However, the task of constructing a hermeneutics of the situation at the programming level would not immediately alleviate the judge's evidentiary job. The interdisciplinary challenges of translation noted in Section II would still be present, and they probably require additional technological expertise in order to gauge the narratives told in court by the parties involved.[66]

Issues are also raised by a focus on legal responsibility for AI, because per Mary Midgley, what "actually happens to us will surely still be determined by human choices. Not even the most admirable machines can make better choices than the people who are supposed to be programming them."[67] This issue arises even in inquiries into negligence and *dolus eventualis*, because while[68]

> humans may classify other drivers as cautious, reckless, good, and impatient, for example, driverless cars may eschew discrete categories … in favor of tracking the observed behavior of every single car ever encountered, with that data then uploaded and shared online – participating in the collective development of a profile for every car and driver far in excess of anything humanly or conceptually graspable.

---

[61] Ryan Abbott & Alex Sarch, "Punishing Artificial Intelligence: Legal Fiction or Science Fiction" (2019) 53:1 *UC Davis Law Review* 323 ["Punishing Artificial Intelligence"] at 348–349 (emphasis in the original).
[62] Ibid. at 355.
[63] Regarding corporate liability, see Chapter 4 in this volume.
[64] Regarding programmer liability, see Chapter 2 in this volume.
[65] "Punishing Artificial Intelligence", note 61 above, at 354.
[66] Regarding evidentiary issues raised by robot testimony, see Chapter 8 in this volume.
[67] Mary Midgley, *What Is Philosophy for?* (London, UK: Bloomsbury Academic, 2018) at 207–208.
[68] Brian Cantwell Smith, *The Promise of Artificial Intelligence: Reckoning and Judgment* (Cambridge, MA: The MIT Press, 2019) at 60.

This chapter argues that human agency matters at all levels of evaluating an ADS. Abbott and Sarch assert that[69]

> [o]ne conceivable way to argue that an AI (say, an autonomous vehicle) had the intention (purpose) to cause an outcome (to harm a pedestrian) would be to ask whether the AI was guiding its behavior so as to make this outcome more likely (relative to its background probability of occurring). Is the AI monitoring conditions around it to identify ways to make this outcome more likely? Is the AI then disposed to make these behavioral adjustments to make the outcome more likely (either as a goal in itself, or as a means to accomplishing another goal)? If so, then the AI plausibility may be said to have the purpose of causing that outcome.

However, humans create AI programmes. The potential to programme ADS in a certain way, and the decision of whether to do that or not, brings us back to the case of the trolley discussed in Section I, and it supports the position that human agency is relevant to evaluating ADS. Another way of considering the role of humans in ADS is provided by what Philippa Foot calls the "doctrine of the double effect," "the distinction between what a man foresees as a result of his voluntary action and what, in the strict sense, he intends"; in other words, he "intends in the strictest sense both those things that he aims at as ends and those that he aims at as means to his ends."[70] Per Foot, the thesis is that it is "sometimes permissible to bring about by oblique intention what one may not directly intend."[71] But can a human inside an ADS exercise free will when it comes to the vehicle's actions?

Could we turn the tables on an ADS, and say that in the current state-of-the-art there is always the abstract risk that such vehicles will swerve out of the control of its human driver, on account of its newly developed intent or other basis, and that because the human driver is unable to anticipate such actions in a preventable way,[72] the risk is agent-relative to the manufacturer-engineer-designer and should be allocated solely to them, i.e., Abbott and Sarch's first solution?[73] This would avoid the

---

[69] "Punishing Artificial Intelligence", note 61 above, at 358.
[70] Philippa Foot, "The Problem of Abortion and the Doctrine of the Double Effect" (1967) 5 *Oxford Review* 1 at 1.
[71] Ibid. at 2.
[72] Regarding the concept of trust in medical robots, see Chapter 3 in this volume.
[73] For the terms "agent-relative" and "agent-independent," see Peter Westen, "The Ontological Problem of 'Risk' and 'Endangerment' in Criminal Law" in R. Antony Duff & Stuart Green (eds.), *Philosophical Foundations of Criminal Law* (Oxford Scholarship Online, 2011) 304 at 306.

question of whether ADS can act intentionally in criminal law, as the risk would be independent of the mental state of the human driver. Depending on the jurisdiction, it may also bring back questions of legal personhood regarding corporate entities.

If the focus of liability is on the manufacturer-engineer-designer, how should liability be understood if an ADS device containing algorithms thinks for itself and gains a certain autonomy? Mary Shelley's fictive monster constructed by Victor Frankenstein began to think for itself. How would a manufacturer-engineer-designer liability for future actions not included in its original programming be understood, e.g., when the machine learning is unsupervised? If we want to distribute risk evenly, we would probably need empirical research to do the math regarding the probability of harm in terms of percentages. For the legislator, the need for refined probabilities of risk could mean an increase in highly refined regulatory offenses. This approach would require a novel definition – or should we say concept? – of conduct, depending on whether there is any active role left for the human driver-passenger. In narratological terms, the driver finds herself in an inbuilt plot of a technological narrative from which she cannot escape; she cannot constrain the non-human actant other than by trying to take over the system when she sees something go wrong, and only if she sees it in time. Thinking about ADS in this way would mean that many advantages of the automatic part of automatic driving systems are done away with, and yet the driver still constantly faces the risk of a future criminal charge.

V   Conclusion: The Outward and Inward Appearances of Intention

This chapter argues for the development of a hermeneutics of the situation to address the issues raised by ADS. As surveyed in the chapter, the issues are many. The *factum probandum* with regard to foresight and the dispositional element included in the concept of *dolus eventualis* are surrounded by challenges. In accidents involving ADS, the debates regarding what the evidence shows in concrete cases will be massive. How is one to decide that a specific human or non-human defendant's disposition suffices for a conviction? These legal determinations will require a careful distinction between the outward appearance, i.e., apparently careless driving, and the legal carelessness of the driver, i.e., his or her indifference to the outcome. The externally ascertainable aspects of any defendant's

action must be taken into consideration in order to make a coherent finding on the elements "knowingly and willingly" of intent.

Some final examples illustrate the importance of the distinction between outward appearance and inward intent or carelessness. Intelligent Traffic Light Control systems can perceive traffic density by means of floating car data apps, which then decide who gets right of way; they are based on the algorithmic ideal of the traffic light talking back to the vehicle. Numerous cases of ADS spontaneously braking in situations where traffic did not require it have occurred, merely because the autopilot thought it recognized the location as one where it had braked earlier. This ADS response is literally a hermeneutics of the situation, but technically a fake negative, in which the human involved may suffer the consequences. In a 2019 Dutch criminal case, the defendant's vehicle had swerved from its lane and collided head-on with an oncoming car. Based on Article 6 of the Road Traffic Act, the defendant was subject to the primary charge of culpable behavior in that he caused a traffic accident by his recklessness, or at a minimum the subsidiary charge that he caused the accident by his considerably careless and/or inattentive behavior, and as a result a person was killed.[74] The defendant pleaded not guilty, arguing that the threshold test for recklessness and/or carelessness had not been met, as he had taken his eye off the road for only a few seconds because he had assumed that the Autosteer System of his Tesla was activated. This position was not given any weight by the court. The defendant was found guilty because his lawyer admitted his client had taken his eye off the road for four to five seconds, and this action was characterized as "considerable" inattentiveness.[75] In the well-known *Vasquez* case in the United States, an investigation by the National Transportation Board suggested that the driver had been visually distracted. Generally speaking, distraction is "a

---

[74] All Dutch laws can be found at www.Overheid.nl, "Law Bank," www.wetten.nl. In 2013, the then Dutch Ministry of Transport, Public Works and Water Management published an English translation of relevant sections of both codes. Article 6 reads: "All participants in traffic are forbidden to behave in such a way that a traffic accident attributable to them occurs in which another person is killed or sustains serious physical injury or physical injury such that temporary illness occurs or that person is prevented from engaging in normal activity." The translation "attributable to them" does not capture the essence of the Dutch text, which refers to the doctrinal *culpa* in the sense of fault rather than criminal intention, so what is meant is attribution in the sense of culpability.

[75] Dutch District Court Oost-Brabant, Decision of September 3, 2019, 01/860055-19, ECLI:NL:RBOBR:2019:5057.

typical effect of automation complacency,"[76] and it suggests the need for driver training. But in this case, the driver had presumably been gazing downward to the bottom of the center console for 34 percent of the time that the ADS was moving, 31.5 minutes, and about "6 seconds before the crash, she redirected her gaze downward, where it remained until about 1 second before the crash," so that there was no time to react and avoid the crash.[77] The driver had supposedly been streaming a television show on her mobile phone during the entire trip.[78] The vehicle "was designed to operate in autonomous mode only on pre-mapped, designated routes."[79] Did the fact that it was a test drive, and a short one at that, on a test road, make the driver behave irresponsibly by watching television while driving? Technical issues with regard to the vehicle and/ or the company's instruction of its employees aside, any driver of a non-automatic vehicle who acts in this way will probably be held criminally responsible, at the very least for behaving negligently. The difference between a traditional driver and a human operator of an ADS has not made great differences in court verdicts yet, in part because inattentiveness attracts liability of some sort. It is, after all, always a human driver who sets the ADS into motion.

Precisely because it is a mental phenomenon, the general concept of intent, as Ferry de Jong contends, is "an essentially 'normative' phenomenon."[80] It "designates … a criminally relevant manifestation of intentional directedness *between* a subject and the social-life world," so that "this intention externalizes itself in the action performed and is thereby rendered amenable to interpretation," which as a "rule-guided process consists of a pre-eminently *hermeneutic* activity: by way of outward indications, the internal world of intentions and perceptions … is reconstructed."[81] If the liability of ADS is to be hermeneutically ascertained, compared to being explained by means of, e.g., statistical evidence on traffic accidents in specific locations that invite some people's dangerous

---

[76] See US, National Transport Safety Board, Highway Accident Report: Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian (Washington, DC: National Transport Safety Board, 2018) at section 1, www.ntsb.gov/ investigations/AccidentReports/Reports/HAR1903.pdf.

[77] Ibid. at 18 and 43.

[78] Ibid. at 24.

[79] Ibid. at 8.

[80] Ferry de Jong, "Theorizing Criminal Intent: A Methodological Account" (2011) 7:1 *Utrecht Law Review* 1 at 1, https://utrechtlawreview.org/articles/10.18352/ulr.144.

[81] Ibid., emphasis in the original.

driving, a hermeneutics of the situation in at least two forms is required. First, in court surroundings, the situation would include the doctrinal, conceptual situation of a specific case, a "hermeneutics of the [legal] signification,"[82] a thorough investigation of the defendant's acts and omissions, and the situation of technology in the sense of the state-of-the-art of the vehicle involved. Second, on the meta-level, such hermeneutics would include a debate on the acceptance of various forms of criminal liability in relation to forms of legal personhood, its technological thresholds and machine autonomy, and societal views on the subject.

A hermeneutics of the situation for ADS is necessarily interdisciplinary. The humanities can contribute to the construction of a hermeneutics of the situation partly by means of narratological insights, because insight is needed into the analysis of narratives, both as story, the what, and discourse, the how, in the pre-trial phase and in court, as well as on the narrative structure of technological proposals and their underlying arguments. As long as technological devices are not fully predictable, explanation must be complemented by understanding. To the French philosopher Paul Ricoeur, "narrative is 'imitation of action' (*mimesis*),"[83] which means that "to say *what* an action is, is to say *why* it is done."[84] In legal surroundings, narratives of judgment therefore address intent and legal imputation. The humanities can also contribute to a hermeneutics of the situation because the technological context of ADS raises the ethics of programming. There is good reason to add a legal-hermeneutic methodology of understanding when deciding ADS cases, lest our technological "swerve" swerves out of control, and we gain no further knowledge of causes and the secret motions of things as Bacon urged us to.[85]

---

[82] "End of Doctrine", note 3 above.

[83] Charles Reagan, "Interview with Paul Ricoeur" in *Paul Ricoeur: His Life and His Work* (Chicago, IL: University of Chicago Press, 1996) 75.

[84] Paul Ricoeur, *Oneself as Another*, translated by Kathleen Blamey (Chicago, IL: University of Chicago Press, 1995) at 63 (emphasis added).

[85] Stephen Greenblatt, *The Swerve: How the Renaissance Began* (London, UK: Vintage Books, 2012) at 7, "swerve" being "an unexpected, unpredictable movement of matter," coined by Lucretius in Lucretius, *De Rerum Natura* (On the Nature of Things), translated by John Watson (London, UK: Bell & Daldy, 1870).