INTRODUCTION

What Is an Autonomous Vehicle?

Two contemporary art works serve as starting points for this book about the material and cultural politics of the emergence of the driverless car as an autonomous, moral **agent**. The first is *Proof of Stake–Technological Claims*, an exhibition in Kunstverein, Hamburg, Germany, curated by Simon Denny and Bettina Steinbrügge (2021). The second is *Ethical Things* by Simone Rebaudengo and Matthieu Cherubini (2016), which is a playful commentary on machine **ethics**, the niche area of inquiry focused on developing detailed scripts for computers to make complex moral decisions. *Proof of Stake–Technological Claims*, 'questions how the framing of processes or objects as "technological" performs cultural work, how they become embedded in our institutions, our social lives, and our thinking.'[10] The exhibition brings together a variety of mixed media artworks that evoke the intricate and powerful ways that digital artefacts and social visions and values interact to transform, or re-organise, social and cultural life.[11] In the book accompanying this exhibition, scholars were invited to write about how media objects organise everyday life, and Simon Denny created NFTs to accompany every object featured in the collection. Denny created the image of a Roomba to go alongside my essay titled 'What is an Autonomous Vehicle?' making a tongue-in-cheek statement that prompts deeper analysis of what exactly an autonomous vehicle is.

The Roomba was built on theoretical advances, research and inventions by Rodney Brooks at MIT, who argued that the human brain was not the most apt metaphor for **AI**. He instead emphasised interaction, situation, and embodiment.[12] The traditional approach to AI and robotics at the time, the 1980s-1990s, was to formally program computer systems to solve problems by following sets of rules from the top down. This was long before the era of big data and **machine learning**. Observing cockroaches led Brooks to question how such small creatures can navigate complex, unfamiliar terrain. The answer lies in their ability to develop internal maps based on their own experiences of navigating a space, rather than relying on a pre-existing, top-down representation. Brooks decided that a robot would not have a centralised representation of the world but would slowly build up its own internal cognitive map of the world it was encountering.

So, he broke from programming an AI system with a top-down representation of the world and gave it the tools to form its own representations to learn about and navigate the world. He convinced much of the robotics community that the human mind should no longer be the blueprint for robotics, but that the capacity to cope with the real world should be instead. Robotics should follow the line of evolutionary complexity, he argued, and only pursue the modelling of human intelligence once animal intelligence is achieved.[13] This work resulted in the invention of the Roomba, the 'autonomous' vacuum cleaner, as well as its parent company, iRobot. A Roomba is made for a space that is free from obstacles—it does, however, know how to go around them to some extent—and where everything is at right angles, with only the lowest of raised thresholds between rooms.

agent
pg.58

ethics
pg.63

AI
pg.60

machine
learning
pg.64

10 See more here: https://www.e-flux.com/announcements/366536/proof-of-stake-technological-claims/.

11 'Proof of stake' is related to 'proof of work,' a protocol adopted by cryptocurrency applications as evidence of computational labour to verify that a transaction has occurred and is appended to the blockchain. Proof of stake is a parallel protocol that addresses the excessive energy consumption required by the PoW protocol. Instead, it relies on verification based on the amount of cryptocurrency owned.

12 J. Bruder, *Cognitive Code: Post-Anthropocentric Intelligence and the Infrastructural Brain* (Montreal: McGill-Queen's University Press, 2020), pp. 4-5.

13 R. Brooks, 'New Approaches to Robotics,' *Science* 253, no. 5025 (1991): pp. 1227-1232. https://doi.org/10.1126/science.253.5025.1227.

As a Roomba owner, I know that it cannot navigate the 1920s layout and flooring of my Berlin apartment. In other words, much like the driverless car, the Roomba works when the world around it transforms to accommodate its limits and expand its potentials. Choosing the image of a Roomba rather than, say, KITT from Knight Rider, Denny is suggesting the latter is not much more than a *tool*—and a *domestic* one at that. Very quickly, Denny's image muddies how we might think about what an autonomous vehicle is if its ability to manoeuvre through space comes not from magic, or from 'on top' programmed in response to rules, as we shall see, but through the observations and recognition of objects around it; and by building maps that are not associated with *place* as humans might create and know a place, but as territories to be navigated. Where the Roomba is concerned, these territories are also intimate spaces. In 2022, iRobot, Roomba's parent company, was bought by Amazon for US$1.7 billion, prompting concerns that Amazon would now have granular information about the layout of people's homes, and that that data could feed marketing and advertising.[14] Tesla's driverless cars also quite literally rely on embodied intelligence and mimic how insects, such as ants, share knowledge: information about a territory that one car captures is shared across entire fleets of cars. Ants do it through pheromonal exchange; Teslas use data exchange via networked infrastructure known as 'the cloud.'

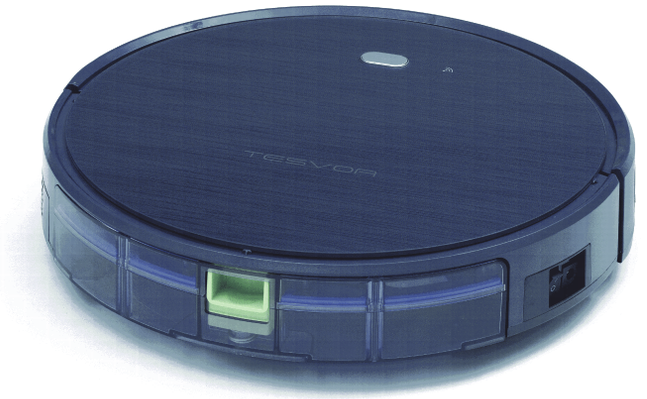

Image 1: 'Autonomous vehicle' by Simon Denny

iRobot sold its technology to the US military to develop autonomous weapons systems. The anthropologist Lucy Suchman comments on the outsourcing of human 'dirty work' to machines: first the work of cleaning, and then killing at a distance.[15] The driverless car as a kind of autonomous vehicle grapples with what it means to outsource difficult moral questions to machines.[16] This book is about a tension inherent to many applications of AI and algorithmic technologies that are mirrored in the case of the driverless car. Are ethics and responsibility matters of values programmed into computational design; of human social, organisational,

14 Alex Webb, 'Amazon's iRobot Deal is About Roomba's Data Collection,' *Bloomberg*, August 5, 2022, https://www.bloomberg.com/news/articles/2022-08-05/amazon-s-irobot-deal-is-about-roomba-s-data-collection?leadSource=uverify%20wall.

15 L. Suchman, 'Situational Awareness: Deadly Bioconvergence at the Boundaries of Bodies and Machines,' *Media Tropes* 5, no. 1 (2015): pp. 15-16.

16 *Reuters*, 'Tesla Faces U.S. Criminal Probe Over Self-Driving Claims: Sources,' *Reuters*, 2022, https://www.reuters.com/legal/exclusive-tesla-faces-us-criminal-probe-over-self-driving-claims-sources-2022-10-26/.

and design practices of implementing values in the imagination and building of technology; or some combination of the two? In popular and academic discussions, a driverless car is imagined as an autonomous agent tootling along, making ethical decisions internally, computationally, in response to external situations such as unexpected crashes, and justifying why it did so. This is what James Moor refers to as an explicit ethical agent, a seductive idea. Autonomous driving is imagined as a human driver replaced by a machine driver, very much in the way that AI is imagined as proxy for the human, and something eventually exceeding it. Is the autonomous vehicle a 'connected and networked car' that is part of a larger data assembly, or the replacement of the human driver, or a robot that is independent? These imaginaries and fantasies of an AI co-exist in the case of this technology as they do for many others.

However, driving is not just about human drivers operating a car; as this book discusses, automobility is a sprawling, infrastructural, human enterprise.[17] It includes public transport systems, city planning, the energy industries, industrial automotive engineering and design, human mobility, and insurance industries amongst others. Such an infrastructural view is hard to advance in the context of the driverless car, given the persistent discursive and rhetorical mimesis of artificial intelligence as human intelligence; and of the human as an individual rather than as constituted by multiple social relationships and identities. This tension between distributed and atomised perspectives on technologies like AI and in the case of the driverless car play out in this book through the study of crashes and autonomous driving infrastructures; and the ethics of autonomous driving, a popular and short-lived discourse about how values could inform the design of an ethical artefact.

Central to the ethics of autonomous driving is the Trolley Problem, a thought experiment in Philosophy; the Moral Machine project, an online game that crowdsources ethical decisions, and how these focused attempts at machine ethics sit alongside the infrastructural complexity of the emerging driverless car. I argue that statistical analysis performed on a vast scale and with great complexity, will replace human moral conflict resolution. We are witnessing and living through a shift in how human moral and philosophical inquiry is becoming entwined with data science and the capabilities of AI systems.

17 J. Urry, 'The "System" of Automobility,' *Theory, Culture & Society* 21 (2004): pp. 25–39. https://doi.org/10.1177/0263276404046059.

The Trolley Problem and the Moral Machine

The Trolley Problem revolves around the doctrine of double effect, exploring the difference between *killing* and *letting die*. It was originally intended to provoke thinking about women's rights over their bodies. It goes like this:

> Suppose you are the driver of a trolley. The trolley rounds a bend, and there come into view ahead five track workmen, who have been repairing the track. The track goes through a bit of a valley at that point, and the sides are steep, so you must stop the trolley if you are to avoid running the five men down. You step on the brakes, but alas they don't work. Now you suddenly see a spur of track leading off to the right. You can turn the trolley onto it, and thus save the five men on the straight track ahead. Unfortunately, Mrs. [Philippa] Foot has arranged that there is one track workman on that spur of track. He can no more get off the track in time than the five can, so you will kill him if you turn the trolley onto him. Is it morally permissible for you to turn the trolley?[18]

This book discusses this in detail, examining how it became a provocation for engineers and manufacturers to consider a future with autonomous vehicles (AV) that will make decisions independently—i.e., 'autonomously'—and without human intervention. It provokes discussion on the decisions that a fully autonomous vehicle's AI software will need to compute in the event of a crash, particularly when there is potential for loss of human life, harm to humans, damage to property, or both.[19] The Trolley Problem became shorthand for a public and academic discussion about how to 'assess the need for a moral component to automated vehicle decision making during unavoidable crashes, and to identify the most promising strategies from the field of machine ethics for application in road vehicle automation.'[20] It was predicted to be one of the first frames for 'artificial morality' because driverless trains and cars would be the first 'robot technologies' we would encounter in everyday life.[21]

The AI scientist Gary Marcus wrote for the *New Yorker* magazine about the development of driverless cars in 2012:

> That moment will be significant not just because it will signal the end of one more human niche, but because it will signal the beginning of another: the era in which it will no longer be optional for machines to have ethical systems.[22]

Marcus does not tell us how this will happen, or what kinds of ethics and ethical systems we will have, but in the years since, many actors are staking a claim to exactly this. 'AI Ethics' has become a bustling area of inquiry and practice across law and regulation, the technology industry, academia, and civil society. The Trolley Problem has generated sensationalist media attention and academic scholarship thanks to scholars who have positioned it in both contexts.

The problem prompted the question can and should a driverless car be programmed to make moral decisions? What kinds of values would inform this programming? We certainly aim to advance scholarship and

18 J.J. Thomson, 'The Trolley Problem,' *The Yale Law Journal* 94 (1985): p. 1395.

19 P. Lin, 'The Ethics of Autonomous Cars,' *The Atlantic*, October 8, 2013. https://www.theatlantic.com/technology/archive/2013/10/the-ethics-of-autonomous-cars/280360/.

20 N. Goodall, 'Ethical Decision Making During Automated Vehicle Crashes,' *Transportation Research Record: Journal of the Transportation Research Board* 2424 (2014).

21 W. Wallach and C. Allen, *Moral Machines: Teaching Robots Right from Wrong* (New York: Oxford University Press, 2009), pp. 13-14.

22 G. Marcus, 'Moral Machines,' *The New Yorker*, November 24, 2012. https://www.newyorker.com/news/news-desk/moral-machines.

practice concerning values in the design of technologies. However, a major challenge today is deciding where this should occur. Should it happen upstream in places like Silicon Valley or Shanghai, where technologies are designed and conceived? Should it be programmed directly into the artefact itself? Should it be something the technology 'learns' in response to its environment, like Brooks' cockroach or Roomba? Should it be embedded within the infrastructural elements that constitute the technology, such as its data? Or should it be within the techno-managerial systems that will develop around and validate the technology, such as laws, audits, checklists and model cards? Clearly, it must happen at all these levels. However, the focus on the ethics of autonomous driving through the lens of the Trolley Problem has narrowed the academic and scientific discussion about ethics and technology. It has done so by promoting the idea that a driverless car can be programmed to make fine-grained moral decisions.

While the Trolley Problem reached the mainstream media and TED Talks, academics came up with proposals for machine ethics as detailed in this book. The most well-known of these is the Moral Machine, a project from MIT Media Lab. This is an online 'serious game' that invited the public from around the world to respond to thirteen scenarios loosely modelled on the original provocation of the Trolley Problem. The Moral Machine assembled a global dataset of moral values about how a driverless car should prioritise different human and nonhuman lives in the case of an unexpected accident.

What is crucial about the Moral Machine is that it shifts the framing of ethical decision-making to statistically modelled governance, driven by the logics of risk, speculation and probabilistic correlation, as this book discusses. Central to this is a shift in the 'ethical' from something humans do to something that regulates the behaviour of a vehicle, and, by extension, the world around it. Hence, this work does *not* follow the normative formulation: 'What is the ethical framework that should inform the decision-making by this technology?' Rather, it identifies the social visions, epistemic, industrial, and technical infrastructures that advance the assumption that gamified thought experiments and statistical calculation of values are the most effective approaches to governing autonomous technologies in complex urban spaces.

Researchers have argued convincingly why the Trolley Problem is an inappropriate frame for autonomous vehicle ethics.[23] They have also argued that the moral issue lies not in the moment of the decision but long before the trolley even reaches the track.[24] Media attention to the Trolley Problem has been largely uncritical and reactive, like associating ethical decision-making with programming an on/off switch, as if on a table-top fan, to decide whom the driverless car should kill. Its media popularity has eventually declined.[25] Yet, the ethics of autonomous driving has sparked research across the fields of computer science, moral and social psychology, behavioural economics, mathematics, and philosophy. This book details some of these approaches.

As I have found in interviews with engineers and philosophers, machine ethics, the Trolley Problem, and the Moral Machine were only intended to provoke thinking about this surface area of ethical concerns, not as directions for how to implement and shape ethical action or decision-making. So, in this book, I take them as sites of inquiry to ask how they collectively constitute the material, social, and cultural infrastructures of autonomous driving. These infrastructures are crucially tied to how we generate and legitimise knowledge. They are supported by technical, scientific, policy and cyber-physical systems knowledge, which extend the limits of human reason, institutions, and capabilities. Additionally, they emphasise the advantages of statistical risk,

23 H. Roff, 'The Folly of Trolleys: Ethical Challenges and Autonomous Vehicles,' *Brookings*, December 7, 2018, https://brookings.edu/research/the-folly-of-trolleys-ethical-challenges-and-autonomous-vehicles/.

24 S. Nyholm and J. Smids, 'The Ethics of Accident-Algorithms for Self-Driving Cars: An Applied Trolley Problem?' *Ethical Theory and Moral Practice* 19, no. 5 (2016): pp. 1275–1289. https://doi.org/10.1007/s10677-016-9745-2.

25 Z. Cassani Davis, 'Would You Pull the Trolley Switch? Does It Matter?' *The Atlantic*, October 9, 2015. https://www.theatlantic.com/technology/archive/2015/10/trolley-problem-history-psychology-morality-driverless-cars/409732/; I. Bogost, 'Enough with the Trolley Problem,' *The Atlantic*, March 30, 2018, https://www.theatlantic.com/technology/archive/2018/03/got-99-problems-but-a-trolley-aint-one/556805/.

data-driven **models**, and AI-ML as organising systems for efficient and effective 'smartness.'[26] What's often missing in this is how cultural symbols, institutions, narratives, and scientific and technological development also shape what 'autonomous' driving means as a social, cultural, and shared urban phenomenon. Further, I am arguing that in bringing autonomous driving into the world, proposals for computational ethics as governance *of* the car and by the car are becoming computational governance *of the world around and outside the car*.

To trace how this is underway I trace figurations of the driverless car in terms of three overlapping cultural ontologies: as car, AI/robot imaginary, and big data infrastructure. This reframing allows for an expanded discussion of the materiality of autonomous driving-in-emergence, and how this materiality shapes how we talk **autonomy** and artificially intelligent cars into being.

models
pg.65

autonomy
pg.63

26 O. Halpern and R. Mitchell, 'The Smartness Mandate: Notes Toward a Critique,' *Grey Room* no. 68 (Summer 2017): pp. 106–129. https://doi.org/10.1162/GREYa00221.

Ethical Things

The second artwork, *Ethical Things* grapples with the implications of computationally guiding moral decision-making through a set of speculative 'smart-ified' objects.[27] One of these objects is a portable, electric, swivel-mounted fan with a dashboard with dials, connected to the internet. It sits on a table between two people. The fan records information entered about the two people to determine which one of them it should turn towards and direct its breeze. Data about the people can be entered in arbitrary terms as far as fans and breeze go: educational levels, weight, religion and gender. If it cannot make an assessment based on the information it has, then it sends the question to an Amazon Mechanical Turk worker (a 'Turker') online. Turkers decide who the fan should turn towards, and they are expected to offer a short justification for their choice.

The results are hilarious, bizarre, and a bit ridiculous; the designers explicitly want the results to appear so. For example, in the video that accompanies the project, we see a response from a Turker that the heavier of the two people should be fanned because fat people sweat more. In some cases, the Turker does not know the answer. We never see the far-flung data worker, and the data worker has no stakes in which way the fan will turn.

27 Read more about the artwork here: http://www.simonerebaudengo.com/project/ethicalthings.

Image 2: The Ethical Fan. Image courtesy of Simone Rebaudengo.

In *Ethical Things*, the designers, Rebaudengo and Cherubini, are responding to the Trolley Problem and the Moral Machine. They are also mounting a critique of a new, niche field, '**value alignment,**' a research topic within the governance of AI and ethics that refers to ensuring that the operational behaviour of an AI system aligns with the values and ethical principles considered important by its human users and stakeholders. Related to 'value alignment' is machine ethics, in which researchers and developers work on methods to program ethical decision-making into AI, such as through ethical theories or rule-based systems. Machine ethics intends to create 'ethical machines' that will adhere to existing human value systems. This is already happening with large transformer models being explicitly programmed to not respond in misogynist, racist, ableist, or other discriminatory ways. We certainly want this.

value alignment pg.69

Yet, whether in large language models (LLMs) now or driverless cars of the future, value alignment and machine ethics must confront the challenge of reconciling different sets of values. They also face the complexity of programming ethics 'into' systems on a scale as large as cities and societies, as if the best ethical decisions were always arrived at through a bulletproof system of reasoning. This might work in narrow or small-scale situations, but in others, it can fail spectacularly. In a recent case of over-correction for discrimination and data bias, Google's Gemini image-generation model produced historically inaccurate images such as a female Pope, English kings with unlikely ethnic and racial ancestry, and Black Nazis.[28] In July 2022, OpenAI developed a technique to automatically 'insert' diversity into image generation technologies,[29] and others, like Google, adopted a similar approach.

Ethical Things is a light-hearted provocation about what it means to design ethics in—i.e, formalise something as 'ethical' at the level of algorithms, data and code. The artists write on their website:

> How can such systems be designed to accommodate the complicatedness of moral and ethical thought processes, especially when human lives are involved? Just like choosing the color of a car, ethics can become a commodified feature in autonomous vehicles (AVs) that one can buy, change, and repurchase, depending on personal taste.[30]

Not all technologies are made ethical at the level of their internal computational functioning. Ethical design can happen in how technology is envisioned, how it is assumed to address social problems, how it inhabits the earth and uses its resources, and so on. This book is about *expanding* the surface area of what and where we think the ethical lies—beyond (but also including) algorithmic actions or thought experiments. So, critically, identifying sites of *where ethics* must happen is a significant preoccupation of the field of 'AI Ethics.' For instance, initiatives towards educating undergraduate computer science students assume that as they eventually enter the workforce, they will integrate more responsible attitudes in their practice.[31] Academics teaching computer science are advancing innovative ways to do this, such as through science fiction literature, with the intention of making the distinction between *how to think* both normatively (what we should do) and descriptively (detailing concerns) about technology and values.[32] As I discuss in the next

28 Benj Edwards, 'Google's hidden AI diversity prompts lead to outcry over historically inaccurate images,' *Ars Technica*, February 22, 2024, https://arstechnica.com/information-technology/2024/02/googles-hidden-ai-diversity-prompts-lead-to-outcry-over-historically-inaccurate-images/.

29 Read more here: https://openai.com/blog/reducing-bias-and-improving-safety-in-dall-e-2.

30 Find the website here: http://automato.farm/portfolio/ethical_things/.

31 Read more here: https://foundation.mozilla.org/en/responsible-computing-challenge/industry-support/.

32 E. Burton, J. Goldsmith and N. Mattei, 'How to Teach Computer Ethics Through Science Fiction,' *Communications of the ACM* 61 (2018): pp. 54-64.

chapter, isolating the 'ethical algorithm' or assigning ethically-minded engineers to specific locations or responsibilities is challenging given the inherently modular structure and design of software work itself[33]; this aside from the corporate incentives that can be at odds with making ethical decisions. In the case of the computational car, I argue that ethics is also distributed across the system. The question of where exactly ethics should be embedded lies beyond the driver. The problem is not 'ethics' but in identifying a site for 'embedding.'

The driverless car has become a site for ethical decision-making, at least speculatively, as indicative of its 'autonomy.' But I argue that the autonomy of autonomous driving is not as we understand *human autonomy*. Autonomous driving is about the transformation of *driving,* rather than the replication of human cognitive capabilities into a computational car. However, the introduction of ethics, particularly machine ethics and value alignment, is confusing here, (mis)leading us to believe that the software in the car is replicating human autonomy to make decisions and have agency. So, I argue that we must focus on the car and mobility in terms of its engineering, human-vehicle interactions, shared urban space, the distinct social imaginaries and ideologies that influence transport and mobility, and the material infrastructure of computing that enable these.

33 D.G. Widder and Dawn Nafus, 'Dislocated Accountabilities in the AI Supply Chain: Modularity and Developers' Notions of Responsibility.' *SAGE Big Data & Society* 10.1 (2023): pp. 1-12.

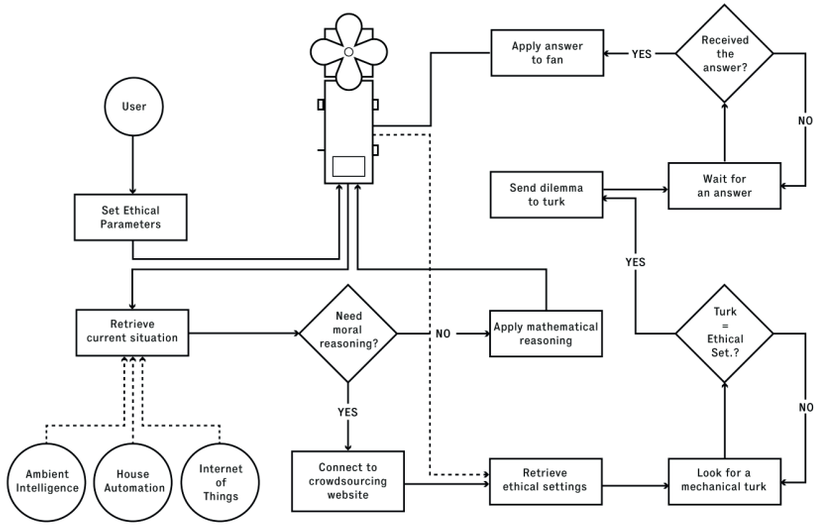

Image 3: Ethical Fan decision tree by Simone Rebaudengo and Matthieu Cherubini

Three Cultural Ontologies[34] of the Driverless Car

Autonomous driving as a techno-solutionist response to poor human driving is not a new concept. Engineers have envisioned it since the 1930s, including designer Norman Bel Geddes, [35] who described how he believed driving would evolve. He stated, 'Everything will be designed by engineering, not by legislation, not in piecemeal fashion, but as a complete job. The two—the car and the road—are both essential to the realisation of automatic safety.'[36] Similarly, in a 1975 oral history interview, the television technology pioneer, Russian-American engineer Vladimir Zworykin, explained his motivation for building the 'autonomous highway' in the late 1950s: 'This growing number of automobiles and people killed in accidents meant something should be done. My idea was that control of automobiles should be done by the road.'[37] Autonomous driving has been in technical development by DARPA and US university researchers since 2004[38] and by Daimler in Germany since the 1980s.[39] However, 'autonomy' in the sense of 'self-driving' is a popular imaginary closely connected to AI, automatons, and robots.

The driverless car is multiple in terms of its constituent cultural ontologies: the future imaginary of the robot car, the connected and networked car-as-big data infrastructure, and the twentieth-century automobile. These ontologies are sustained by and produce new kinds of material and social realities accruing to discourses and meanings about 'autonomous' driving. This allows me to analyse world-making on a broader temporal and situated scale. The AV as AI/robot is considered a replacement for the human driver. The language of autonomous driving echoes the notion of a computational brain inside a vehicular body, or of automation without humans—driverless car; robot taxi; unmanned vehicular systems (unmanned aerial vehicle refers to drones). Back to Denny's image of the vacuum cleaner/autonomous vehicle; he is suggesting its much-vaunted capacity for 'autonomy' as independence, reflection, and decision-making is a red herring. The United States National Highway Transport Safety Authority (NHTSA) might agree. While the language used by Tesla and other manufacturers often slips between self-driving rather than autonomous, the regulators are probing Tesla's advertising of its 'full self-driving mode' because it gives the appearance of independent capabilities when it requires human attention and hands on the wheel.[40] 'Self'-driving suggests the vehicle might have a sense of *self*; or that humans see it as having a self because it can navigate itself. Yet, this 'brain' is actually a distributed configuration of big data infrastructures of cloud connectivity, 3D maps, LIDAR, cameras, running various computational systems. And this is also an automobile that is a twentieth-century media artefact. The driverless car is still a car emerging from a distinct social and cultural history that has significantly shaped twentieth century social, national and economic life. In that sense, the AV of the future is a steady transformation of the contemporary automobile, which has itself been the site of increasing degrees of computational automation over the past 40 years, also known as 'automated driver assistance systems.'[41]

34 'Ontology' is a term from philosophy that refers to the study of the nature of reality and being. In computer science, ontology refers to a structured framework that defines and organises information. It helps describe the relationships between different concepts within a specific domain, enabling computers to understand and process that information more effectively. In anthropology, 'culture' and 'ontology' are used to work through differences in versions of reality that groups of people hold. Here, 'cultural ontology' refers to each aspect of the driverless car as automobile, robot imaginary, and data infrastructure as constituted by different social histories, material infrastructures, and symbolisms. It also refers to shared and popular perceptions in distinct places this research was conducted in.

35 Norman Bel Geddes was a US designer who is significant in this context because one of his best-remembered designs was General Motors-sponsored Futurama exhibit at the New York World's Fair (1939–40). 'Futurama' was important because it was a master plan for the United States twenty years into the future (1959–1960), imagined in terms of automobility and suburban living. Bel Geddes was a significant influence on the automobile and cities. Geddes developed the idea of 'magic motorways' as central to the modern American social life.

36 Quoted in Seiler 2008, Loc 1709.

37 In E. Ackerman, 'Self-Driving Cars Were Just Around the Corner—in 1960,' *IEEE Spectrum: Technology, Engineering, and Science News*, 2016. https://spectrum.ieee.org/selfdriving-cars-were-just-around-the-cornerin-1960.

38 *Defense Advanced Research Projects Agency, Creating Breakthrough Technologies and Capabilities for National Security* (DARPA, 2007). https://archive.darpa.mil/grandchallenge/overview.html.

39 Daimler's 'networked mobility' project, Prometheus, was piloted in 1986. Read more here: https://media.daimler.com/marsMediaSite/en/instance/ko/The-PROMETHEUS-project-launched-in-1986-Pioneering-autonomous-driving.xhtml?oid=13744534.

40 Reuters 2022.

41 P. Leonardi, 'From Road to Lab to Math: The Co-evolution of Technological, Regulatory, and Organizational Innovations for Automotive Crash Testing,' *Social Studies of Science* 40, no. 2

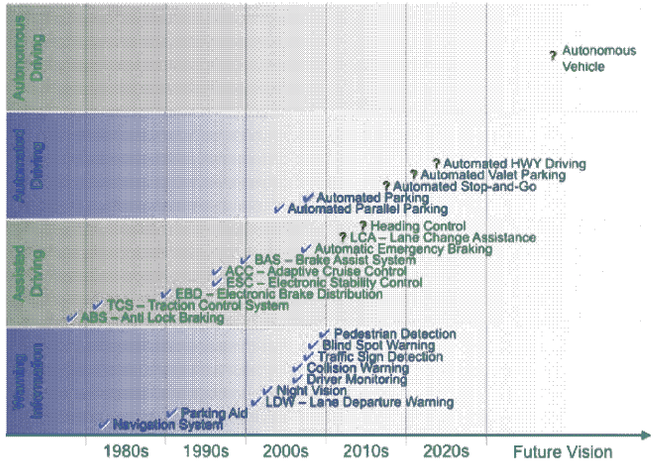

Image 4: Classification and history of driver assistance systems. Sven Beiker's depiction based on sources from Wikipedia. Reproduced with the kind permission of Sven Beiker.

There is an argument to be made for 'a new ontological category for robots somewhere between object and agent.'[42] What I believe this means is to identify how autonomous systems, like embodied robots, are more than constituent data infrastructures, mechanics, and computation, and a little less than a fully explicitly independent agent. This is evoked by design theorist Benjamin Bratton in *The Stack*, which charts the condition of 'planetary-scale computation':

> [T]he integrated design of driverless cars includes navigation interfaces, computationally intensive and environmentally aware rolling hardware, and street systems that can stage the network effects of hundreds of thousands of speeding robots at once...We see not one totality but the production of multiple and incongruous totalities, some of which are 'interfacial regimes' ... may also displace existing geographies.[43]

Hence, critical scholars advocate for an analytical approach that recognises the increasing complexity of these artefacts and their 'incongruous totalities.' They argue that singular narratives are likely to be insufficient for examining the discursive and epistemological roles these artefacts play. It might seem rather obvious that a single thing is not just one thing but is made up of multiple other things depending on who is doing the analysis.[44] However, it is important to emphasise that when AI is presented as having a single historical lineage or future manifestation is misleading. In fact, all computation is a dense, complex assembly of historical systems of systems. *This* is the work of the discursive: to supply the frameworks and representations that gather and hold these different densities together, and tie up them in seamless representations. It is about examining how these different ontologies come together, identifying the 'gnashing and grating juxtapositions' between them, and the 'peculiar new spaces, normal enclaves' that are generated, which 'deliberately reorganise the world.'[45]

(2010): pp. 243–274; K. Bengler, K. Dietmayer, B. Färber, M. Maurer, C. Stiller, and H. Winner, 'Three Decades of Driver Assistance Systems: Review and Future Perspectives,' *IEEE Intelligent Transportation Systems Magazine* 6, no. 4 (2014): pp. 6–22; S. Beiker, 'Deployment Scenarios for Vehicles with Higher-Order Automation,' in *Autonomous Driving*, ed. M. Maurer, J. Gerdes, B. Lenz, and H. Winner (Berlin and Heidelberg: Springer, 2016), pp. 193–211. https://doi.org/10.1007/978-3-662-48847-8_10.

42 R. Calo, 'Robotics and the Lessons of Cyberlaw,' *California Law Review* 513 (2015): p. 119.

43 Bratton 2015, p. 12.

44 This is inspired by the work of the anthropologist, Ann Marie Mol, in *The Body Multiple*, which details how a single diagnosis manifests very differently across the range of health care providers engaged in its management.

45 Ibid.

I take inspiration from a tongue-in-cheek tweet about AI from Ryan Calo, a professor of law, who asks: What is AI? This question is deceptive because it opens up a variety of possible answers indicating its deep technological, social and situated epistemologies and histories. Calo suggests that we are dealing with multiple imaginaries, literacies, discursive anchors, expertise, and epistemic authority figures engaged in the shaping and emergence of AI. And the answer varies depending on whoever is interested in the question, 'what is AI?' He suggests that there is no single object that is 'AI.' To assume that it refers to a synthetic super-intelligence, or a fembot, is to concede to a particular and specific imaginary; these are just one kind of cultural-ontological manifestation of what we refer to as AI. Just as, perhaps, there is no single entity that we might refer to as *the* driverless car.

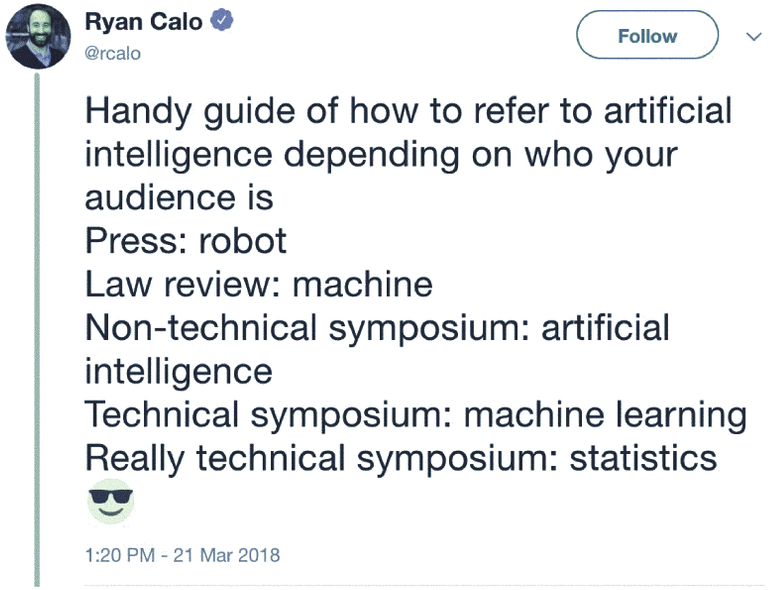

Image 5: Screenshot of a tweet by Ryan Calo, March 21, 2018.

The Driverless Car as the AI/Robot Imaginary

A variety of terms are currently in use across popular and tech media, academic and grey literature, policy documents, and everyday speech to refer to the driverless car. These include robot car, self-driving car, semi- and fully-autonomous vehicle, driverless car, and, more recently, 'connected cars.' None of them are perfect, which is part of the problem. Naming new and emergent technologies can be notoriously difficult in any age. The emergence of the word 'automobile' in the late 1800s was the subject of scornful discussion in the popular press at the time.[46] Using the words 'autonomous' and 'driverless' allow us to imagine that the car is moving 'on its own.'

Autonomous vehicles, broadly understood, already exist in restricted contexts, such as in drones, deep sea exploration, in war and conflict situations, and mobile delivery in controlled environments. The fully autonomous vehicle, with an emphasis on *fully*, does not exist anywhere but in the future. What is the relationship between these past fictions and our present? Is the AV materialising now because its mention in science fiction served as inspirations for scientists and engineers?[47] However, the emergent AV is not a prototype anymore but is an actual artefact being publicly tested.

The role of fiction, imaginaries and metaphors in shaping discourse cannot be overstated. In the movie *Minority Report* (Steven Spielberg, 2002), a 'fully' autonomous vehicle plays a key role at one of many decisive moments. The protagonist, a detective in the 'Pre-Crime' division, John Anderton, is on the run and in a driverless car–which identifies him and locates him on the network. When he realises he has to run, a chase scene ensues and Spielberg dazzles us with shots of AVs moving at high speed in dedicated, elevated channels. Yet, they still maintain that satisfying vroom vroom sound, suggesting that even in the future, the car will not lose its 'feel.' What is distinct about *Minority Report* is that the world of 2054 exists in the smallest details, such as the cartoons on the side of the cereal box, the newspaper that updates itself, and personalised advertising that jumps out at the protagonist.[48] Cinema has a powerful role in generating visions of the future—even more so when a powerful cultural influencer like Spielberg assembles scientists and engineers to help him create a future of ubiquitous computing, with screens and digital interfaces everywhere. David Kirby writes that 'diegetic prototypes' are technologies that exist only in fictional world'—what film scholars call the diegesis—but they exist as fully functioning objects in that world,'[49] and thus demonstrate to us, the audience, the 'need, viability and benevolence' of those technologies.[50] Such 'fictions' can be a valuable jumping-off point to create legitimacy for institutional knowledge and naturalise the problems that necessitate particular kinds of products, and the social orders they generate. The dramatic spectacle of a 'robot' car[51] that is making decisions on its own plays on multiple anxieties, chiefly that of a long-held concern about AI/robot technologies exceeding human control, a popular theme in science fiction. It is no surprise, then, that the emergence of driverless cars, like any

46 A. LaFrance, 'What Should We Call Self-Driving Cars?' *The Atlantic*, March 1, 2016, www.theatlantic.com/technology/archive/2016/03/what-should-we-call-self-driving-cars/471711/.

47 'Science fiction does not merely anticipate but actively shapes technological futures through its effect on the collective imagination'; and 'Science fiction visions appear as prototypes for future technological environments' (Dourish and Bell 2014, p. 769).

48 Spielberg took great pains to make the year 2054 seem realistic and true to life. Before the release of the film, he spoke with the film critic Roger Ebert about his process: https://en.wikipedia.org/wiki/Minority_Report_(film).

49 D. Kirby, 'The Future is Now: Diegetic Prototypes and the Role of Popular Films in Generating Real-World Technological Development,' *Social Studies of Science* 40, no. 1 (2010): p. 41.

50 Ibid.

51 There is an extensive Wikipedia page dedicated to robot cars and other AVs in science fiction cinema, television and literature: https://en.wikipedia.org/wiki/Self-driving_car#In_fiction.

new technological innovation, must prioritise safety to gain the trust of various publics and stakeholders—ranging from potential future users to regulators and auto manufacturers—if it is to gain widespread acceptance.

Isaac Asimov's 1942 short story collection *Runaround* features the Three Laws of Robotics[52] that are popular in the discussion of autonomous driving ethics. These laws are a set of rules governing robot-human relations. The First Law: A robot may not injure a human being or, through inaction, allow a human being to come to harm. The Second Law: A robot must obey the orders given it by human beings, except where such orders would conflict with the First Law. And the Third Law: A robot must protect its own existence as long as such protection does not conflict with the First or Second Law. There is also a 'zeroth' law preceding the first that was added later, which relates to 'humanity' rather than the singular human in the first law. In the early days of this research, it was difficult to come across a mainstream media or academic article that did not mention the Laws of Robotics. And while this is fiction, much of our understanding of AI is actually shaped by metaphor, fiction, speculative figurations, and imaginaries. We rely on metaphor to make sense of the uncertain and unfamiliar.[53]

Sheila Jasanoff and Sang-Hyun Kim's concept of 'socio-technical imaginaries' captures the diverse influences on how technology is 'made.' They define these imaginaries as 'collectively held, institutionally stabilized, and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology.'[54]

The 'imaginaries' frame reveals the connective tissue between the interior associations we have with technology and its broader social and political-economic dimensions. The word 'desirable' is of relevance here, suggesting that intentions and visions are also often hidden from us and must be identified through a triangulation of social and political actors, including more discursive and rhetorical ones. Such as, for instance, the question of what constitutes 'the imagination'?

Here, we can point to the powerful place of media and cultural artefacts, narratives, and tropes in shaping visions of technology and the future, particularly when it comes to artificial intelligence technologies.[55] Identifying the autonomous vehicle in terms of the robot imaginary, suggests that how the 'self-driving' car exists in literature and cinema has a place in how visions of the future AV will inform the creation of conditions influencing its actual design and emergence. This is not suggesting a causal or circular link between imagination, discourse, and design, but recognises that future visions also act as self-fulling prophecies. Consider how the engineer's imagination converges with marketing rhetoric of the AV as a replacement driver.

In a public discussion of their work, the CEO of Waymo, the Alphabet-owned self-driving car project, John Krafcik, said that their product is a replacement for a licensed driver. Interestingly, Waymo never talks about the ethics of autonomous driving explicitly, however, the 'replacement of the human driver' that Krafcik refers to, becomes the popular perception of what autonomous driving is. In the tweets, Krafcik unironically frames 'the human driver' in mechanical and computational terms: a mix of hardware and software. Their driverless car,

52 https://en.wikipedia.org/wiki/Three_Laws_of_Robotics.

53 N. Gilman and M. I. Ganesh, 'Making Sense of the Unknown: AI's Metaphors,' in *AI+1: Shaping Our Integrated Future*, ed. The Rockefeller Foundation (New York: The Rockefeller Foundation, 2020).

54 S. Jasanoff, 'Future Imperfect: Science, Technology, and the Imaginations of Modernity,' in *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*, ed. S. Jasanoff and S. H. Kim (Chicago: University of Chicago Press, 2015): pp. 3-4.

55 Cave et al. 2018.

the Waymo Driver, is similar. However, 'the AI part is the hardest,' Krafcik acknowledges. There is an utterly serious way that AI is referred to as the human driver-like element that will animate autonomous driving. On the other hand, this intelligence is, in fact, many technologies, infrastructures and human labour.

Image 6: Screenshots of tweets by John Krafcik, CEO of Waymo, on December 7, 2020.

In a casual conversation, the social scientist of financialisation, Martha Poon, made a throwaway remark that is uncannily astute in bringing together many elements of this multivalent technology. She said, 'The driverless car is imagined as the perfect little neoliberal subject that will tootle along making decisions for itself.'[56] I believe Poon is referring to the imagination of the AV as *Homo economicus*, the human who makes rational and informed decisions, unfettered from struggle, and disconnected from history or affect. The human who is always trying to improve themselves and become smarter and more successful by making the right decisions.

Artificial intelligence is constructed through a fertile and messy exchange of metaphors about human and machine. In fact, AI itself has been developed as a metaphor for thinking and intelligence. Metaphors are powerfully entangled with epistemology, even when they are not accurate and are constitutive of theory particularly in young fields of research. Theoretical psychology, for example, is rife with analogies of humans as computers and vice versa, that, 'computer metaphors have an indispensable role in the formulation and articulation of theoretical positions' about how the human mind works.[57]

Rodney Brooks has been deeply critical of this metaphor, referring to it as an 'intellectual cul de sac' that does not advance AI.[58] The AV is imagined as an artefact that is humanoid in processing capabilities in the same way that AI/ artificial intelligence is, an 'awesome thinking machine' that will make decisions for itself, automatically or, 'autonomously.'[59] The driverless car as AI, a machine that exists independently embodies a fetish of individuality.

56 In personal communication, Brussels, Belgium, January 27, 2017.

57 R. Boyd, 'Metaphor and Theory Change: What Is "Metaphor" a Metaphor For?' in *Metaphor and Thought*, 2nd ed., ed. A. Ortony (Cambridge: Cambridge University Press, 1993): p. 487.

58 Brooks 2012.

59 S. Natale and A. Ballatore, 'Imagining the Thinking Machine: Technological Myths and the Rise of Artificial Intelligence,' *Convergence: The International Journal of Research into New Media Technologies* 1–16 (2017). https://doi.org/10.1177/1354856517715164.

The Driverless Car as a Big Data Infrastructure

The driverless car as a big data infrastructure underscores the intricate technical landscape that humans are deeply intertwined with. What Drive. Ai and BMW refer to as the AV's 'brain' is, in fact, a vast data-infrastructural network spread over multiple commercial, regulatory, legal, and 'cloud geographies.'[60] The material infrastructures within the emergent AV render it as data platform[61] that runs AI technologies like computer vision and automated decision-making based on multiple sources of data processed through machine learning. Metaphors of a computational brain are materialised by a raft of profitable software companies that are building autonomous driving based on AI technologies. The venture capital firm Comet Labs detailed the 263 companies working on driverless car technology. Many of these are small and relatively unknown, such as Actility and Braiq, but produce important components, sitting alongside bigger names like Google and Uber[62] (see image below). This chart is from 2017 and is likely to have changed considerably in the years since it was published.

The AV exists in and as a 'formidable' 'intelligent vehicular grid,' a big data-infrastructural platform. It is comprised of sensors capturing and processing data about the environment, cameras, radar, Lidar, myriad data processing functions, including machine learning, object recognition, tracking, and coordination, mapping and localisation systems, machine-readable road signs, networking and communication architectures including vehicular cloud computing, computer vision, machine-learning based risk and uncertainty assessments, hard and soft telematics, and driving style analysis among others.[63] It also includes sensors, and emotion recognition software to track sleepiness or distraction.

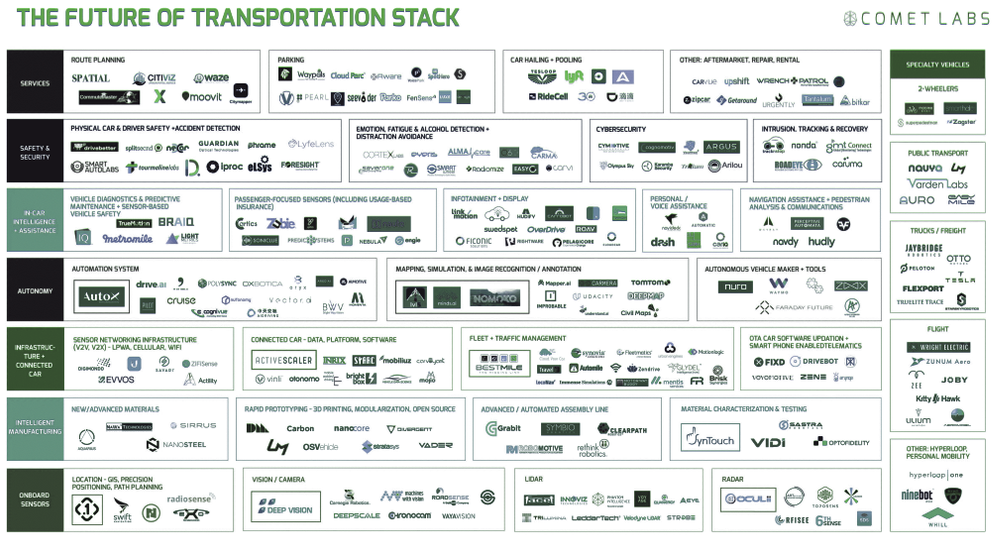

Image 7: **The Future of Transportation Stack by Comet Lab, cited in Stewart (2017) Mapped: The Top 263 Companies Racing Toward Autonomous Cars. Wired, May 10, 2017. https://www.wired.com/2017/05/mapped-top-263-companies-racing-toward-autonomous-cars/**

60 Amoore 2018.

61 L. F. Alvarez León, 'Eyes on the Road: Surveillance Logics in the Autonomous Vehicle Economy,' *Surveillance & Society* 17, no. 1/2 (2019): pp. 198–204. https://ojs.library.queensu.ca/index.php/surveillance-and-society/index.

62 J. Stewart, 'Mapped: The Top 263 Companies Racing Toward Autonomous Cars,' *Wired*, May 10, 2017, https://www.wired.com/2017/05/mapped-top-263-companies-racing-toward-autonomous-cars.

63 M. Gerla, E. K. Lee, G. Pau, and U. Lee, 'Internet of Vehicles: From Intelligent Grid to Autonomous Cars and Vehicular Clouds,' in *IEEE World Forum on Internet of Things (WF-IoT)* (IEEE, 2014), pp. 241–246; E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, 'A Survey of Autonomous Driving: Common Practices and Emerging Technologies,' *IEEE Access* 8 (2020): pp. 58443–58469. https://doi.org/10.1109/9046805.

Thus, the emergent driverless car might be thought of as a computational and data platform, or a 'data assemblage.'[64] Automobility has long been considered hybrid and dynamic, a system of diverse institutional forms from manufacturing and selling automobiles, to highway and 'gasoline delivery infrastructures, traffic rules, parking structures, licensing procedures, and sundry regulatory authorities.'[65] Similarly Dant's 'driver-car assemblage' speaks to the social, cultural, historic, and industrial worlds co-existing and co-evolving.[66] The embodiment of human and automobile fusing have evoked the 'distinctive ontology of the person-thing' like a 'humanized car... or the automobilized person.'[67] Whether we view the driverless car in terms of its software 'brain,' or its automotive-and automobility-infrastructures, or both, we are encountering large-scale relational, complex and distributed social-technical infrastructure. And these are also human systems, comprised of humans in organisations and industries, and humans as users and consumers.

The Driverless Car as a Car

While we have autonomous trucks, autonomous deep-sea exploration robots, and unmanned aerial vehicles (UAVs), the driverless car, or 'self-driving' car is still speculative. The status of 'full autonomy' appears on Automotive Engineering standards as the 'final stage' of automobility. Yet, the driverless car is also just a car, and as such, a very important media technology of the past century. Cars and the open road are symbolic embodiments of independence and autonomy, for example, in the shaping of national identity in the United States during the Cold War,[68] as well as in present-day Saudi Arabia, where women were barred from driving till 2019. In the Hollywood film *Thelma and Louise* (1991), the female protagonists find that freedom exists only by driving off a cliff. The law, as well as the disappointment of relationships with men, can never really be escaped.

Sometimes the merging can be macabre and literal. J.G Ballard's novel *Crash* (1973) drifts from one erotically charged description to another of mangled human bodies and machines fused together in the moment of a car crash. The antagonist, the 'TV scientist,' Vaughan, and his motley crew of car-crash fetishists seek out crashes in-the-making, even causing them, just for the thrill of it. Vaughan's ultimate fantasy is to die in a head-on collision with the actress Elizabeth Taylor. Media theorist Marshall McLuhan refers to the car as an item of clothing: '[T]he car has become an article of dress without which we feel uncertain, unclad, and incomplete in the urban compound.'[69] So, the automobile is a significant media artefact of the twentieth century and this legacy persists in how our cities and everyday urban lives and environments are organised.

The automobile as a social and media technology works in two opposing ways; it individualises and fragments human relations and spaces, but it is also strongly associated with cultures of embodiment and human-machine hybridity. As such, the AV as a car also implies that it is a media object and the site of multiple dynamics of mediation. The first dynamic of individualisation and fragmentation was what Raymond Williams referred to as 'mobile privatisation,' with numerous plays we might make now on the meanings of the words 'mobile' and

64 R. Kitchin and T. Lauriault, 'Digital Data and Data Infrastructures,' in *Digital Geographies*, ed. J. Ash, R. Kitchin, and A. Leszczynski (London: Sage, 2018): pp. 83–94.

65 S. C. Rajan, 'Automobility and the Liberal Disposition,' *The Sociological Review* 54, no. 1 (2006). https://doi.org/10.1111/j.1467-954X.2006.00640.x.

66 Dant 2004, p. 74.

67 Katz 2000, cited in N. Thrift, 'Driving in the City,' *Theory, Culture & Society* 21, nos. 4–5 (2004): p. 47. https://doi.org/10.1177/0263276404046060.

68 Seiler 2008.

69 M. McLuhan, *Understanding Media: The Extensions of Man* (Cambridge, MA: MIT Press, 1964/1994), p. 217.

'privatisation.' Chiefly this refers to the increase in mobility that enabled the (mid twentieth-century North Atlantic and European) home to become more private and self-sufficient.[70] And of course, in tandem with other technologies evolving at the time—from the television set to packaged margarine—that spelled modernity.[71] The notion of mobility was not just literal in terms of car culture and driving, but also includes television and media that offer the outside world as spaces that could be 'travelled to' from within the home. 'Mobile privatisation' also highlighted restrictions on mobility—particularly for older people or those with care responsibilities that kept them at home—cutting them off from wider society. The individual, made private as a result, is cocooned.[72]

Williams was presaging an individualisation and fragmentation that was relatively novel then, and one that we have come to accept as a fact of life now, mediated as our spaces are by mobile phones that create digital bubbles. 'Are we there yet?!' is the annoyed refrain of the traveller who must put up with the tedium of travel and close interactions with other humans as the price paid for the value associated with mobility. Media have rushed in to fill this gap and massage the spaces of friction with other humans; we can shut ourselves off in perceptual and spatial bubbles of headphones and personal screens; we had masks too, during the Covid-19 pandemic. Screens within cars, like video and TV, and our phones, capture our attention—aka data—which translates into algorithmic knowledge about consumer tastes and preferences.[73] This is foreseen as one of the profitable projections of autonomous driving. The opportunities for media and entertainment consumption are thought to explode with the emergence of the fully autonomous vehicle.

70 R. Williams, *Television: Technology and Cultural Form* (London: Routledge, 1974/2003).

71 Owing to the post Second World War butter scarcity, and despite strong resistance from dairy farmers.

72 J. McGuigan, 'Mobile Privatisation and the Neoliberal Self,' *Key Words: A Journal of Cultural Materialism* 11 (2013): pp. 78-79.

73 J. Packer and K. F. Oswald, 'From Windscreen to Widescreen: Screening Technologies and Mobile Communication,' *The Communication Review* 13, no. 4 (2010): p. 314. https://doi.org/10.1080/10714421.2010.525478.

Autonomy and Ethics: Shifting Human Roles in AI Systems

This book is organised around two inter-related cases: the ethics of autonomous driving, and what car crashes reveal about the changing role of the human in AI-systems. The ethics of autonomous driving began as a prompt for engineers to consider the potential negative outcomes of designing and developing a driverless car. However, 'ethics' is not so much about values-informed design but is part of an epistemic infrastructure that advances statistical models of risk assessment and decision-making. That which we believe makes us human—i.e., moral reasoning and working through complex social questions of life and death are now being re-designed as forms of statistical 'reason.' Confusingly, this ability of the computer to 'reason' is conflated with autonomy, another concept that we believe as integral to being human. Thus, the understanding that being human is predicated on achieving specific kinds of cognitive tasks is de-stabilised now that we're building machines that can do similar.

However, there's equally a reinforcing notion that *this* is what being human means. If being human exists in registers and domains of life that are *not* cognitive, then this is not factored into the development of 'intelligent' and 'autonomous' systems. That driving decisions are a cognitive-style activity is another assumption advanced by the ethics of autonomous driving, when, in fact, as many drivers know, we don't think about driving decisions. They become intuitive, learned, and we learn to adapt to situations on the fly. Driverless cars are yet to do this and must learn from prior examples. Hence, the choice of examples—that is, training datasets—matter. The reality of how car crashes occur presents a different perspective. Crashes are inherently systemic rather than individual; crashes reveal the structures that constitute 'driving,' the structures that hum along in the background, becoming invisible in their successful operation. Crashes suggest that driving is an elaborate human-social-organisational and material construction. The speculative

crash imagined by the Trolley Problem and Moral Machine does not engage the realities of roads, test drivers, dashcams, computer vision datasets, testing environments, engineering standards, human operator-car handovers, auto-pilot technologies; or any other material realities implicated in how crashes actually occur.

Aside from the two cases of the shaping of ethics and car crashes, this book has a third preoccupation: how to study the social, material, and cultural dimensions of a significant engineering project like automated driving, its assemblies of interwoven human, built-environmental, cyber-physical, and computational systems. So following scholars of society and technology, and not unlike Denny's 'autonomous vehicle,' I understand materiality to be more than just the exact form or constitution of artefacts, but also how their use changes over time, and what they mean as social and cultural artefacts.

In this process, what are the transformations taking place in how societies are changing, how we govern ourselves, and relate to fellow humans and nonhumans? Our digital technologies are not just literal, material things in themselves; they are made into 'smart,' 'autonomous,' 'intelligent, or 'frictionless' because of the interactions between the social, the individual, the semiotic, and the material. These interactions become invisible at the furthest ends of chains of production, such as users. Projects such as *Anatomy of AI* (Crawford and Joler, 2018) peel back the layers to reveal such processes. *Anatomy of AI* starts with the familiar Amazon speaker device, the Echo, that has a digital assistant function assigned the human name, Alexa, and unpacks the finite natural resources and labour its production requires, and histories of those resource-extractive processes.

There's another element to the material-cultural preoccupation of this book; it is how materiality is also discursive, that the material and cultural forms of technology supply us with the not just *what* we say but *how* we say them, that is, the frames, words, and terms under which we talk about something. As Donna Haraway puts it, 'Discourses are not just "words"; they are material-semiotic practices through which objects of attention and knowing subjects are both constituted.'[74] In this case, the language of the driverless car as autonomous or ethical is derived through identifiable 'epistemic infrastructures'[75] such as the Trolley Problem, Moral Machine, or measures such as **disengagements** the Society of Automotive Engineers' (SAE) Levels of Automation, or McKinsey's various 'pillars of readiness' to rank countries' resources and infrastructures for autonomous driving.

disengagements pg.63

Why examine discourse? Why study its relationship to materiality? Because there is a persistent myth about technology, that technologies like AI are *doing things on their own*, that there is *no way* to arrest how they function, that *autonomy exists* in the car because it is *built that way*. Material-discursivity is the recognition that language, technology use, its form and constitution, its histories, and politics, are all entangled and co-constitutive practices through which that technology is actively made and re-made. There's an aura around technologies like AI as being self-governing, thinking, intelligent, autonomous, and agentic. 'Autonomous driving' and driverless cars are new technologies and so language and narratives must be intentionally made around what they are and might do for us. This ability to name and set in motion ways of thinking about the world and ourselves through technological artefacts is enormously powerful.

However, discursive power is difficult to pin down and is often hidden within technologies; in the sense of being entangled within—in the materiality and making of technology itself. Discourse is also suggestive

74 Haraway 1997.

75 M. Murphy, *The Economization of Life* (Durham, NC: Duke University Press, 2017).

of ideology and values. The work of critical cultural studies of technology such as this one is in part about revealing these material-discursive entanglements. These entanglements often remain hidden because technologies are themselves conveyors of values and beliefs about the world, and embody visions for the future; and by taking them at face value. Whose values and visions of the future does a technology embody and to what extent is this a shared vision that those of us who will partake in it have been consulted on? Values are embedded in design choices, presentation, language, among others, but we are not always adept in discerning these values; and how we analyse hammer, a bridge, or Photoshop are necessarily different. Examples abound: a mobile phone in pink to code femininity; photographic film that renders darker skin as either over-exposed or under-exposed; seatbelts tested and calibrated to taller and bigger humans rather than shorter ones; referring to computational workflows as 'master' and 'slave.'

Interestingly, the 'master-slave' terminology has persisted for decades and was formally retired only in 2020 in the wake of the global protests against the killing of George Floyd by police.[76] Perpetuating discrimination through language only serves to re-commit to the values in that language and by extension, the social world. Studies of biased language in AI and algorithmic systems over the past half decade have been troubling: they are evidence of how deeply our social biases infiltrate the functioning of our machines and by their continued, unchecked use, only further advance those biased values. A now-classic study of bias in word embeddings identified that women's names were more strongly associated with family, or the arts, while men's names were associated with careers and mathematics.[77] Other studies have found similar patterns of discrimination replicated in machine learning systems.

But the question of where and how to stop bias hounds the design and development of AI technologies as we will see, because 'technology' is material and discursive, made-up of environmental, economic, cultural and political aspects; the artefacts and devices carry power, symbolism, and values. So, Denny's Roomba-AV opens up not just how to investigate autonomy but also that we must for it conveys values and intentions that are not evident at first glance. As a study of the material and cultural politics of AI and autonomous technologies, this book assesses how shared social, urban, and cultural life is being re-organised through the vision of such proposals.

76 As this BBC article points out: https://www.bbc.co.uk/news/business-53273923.

77 A. Caliskan et al., 'Semantics Derived Automatically from Language Corpora Contain Human-Like Biases,' *Science* 356, pp. 183-186 (2017). https://doi.org/10.1126/science.aal4230.