Digital navigation and the driving-machine: supervision, calculation, optimization, and recognition

Sam Hind

To cite this article: Sam Hind (2019) Digital navigation and the driving-machine: supervision, calculation, optimization, and recognition, Mobilities, 14:4, 401-417, DOI: 10.1080/17450101.2019.1569581

To link to this article: https://doi.org/10.1080/17450101.2019.1569581

Published online: 06 Feb 2019.

Submit your article to this journal

Article views: 151

View related articles

View Crossmark data

Citing articles: 1 View citing articles
Digital navigation and the driving-machine: supervision, calculation, optimization, and recognition

Sam Hind
Locating Media, University of Siegen, Siegen, Germany

ABSTRACT
In this paper, I explore the navigational implications of a possible driving world. In the last few years, autonomous vehicles (AVs) have garnered significant attention, with much of this scrutiny centered on the technical possibilities, legal restrictions, and utilitarian ethics of AVs. In this paper, I look at how AVs are radically transforming the nature of navigational decision-making. Research into the automation of industrial processes and aircraft fly-by-wire systems suggests that navigational supervision, by humans, will become a significant duty, recalibrating navigation itself. I draw out the implications of automation through three navigational practices of the ‘driving-machine’ I refer to as route-calculation, terrain-optimization, and object-recognition. Attending to these practices assists in the ongoing interrogation of the machinic rendering of automobile space.

ARTICLE HISTORY
Received 15 May 2018
Accepted 21 December 2018

KEYWORDS
Automation; autonomous vehicles (AVs); calculation; driving; navigation; terrain; sensing

Introduction
In the last few years, autonomous vehicles (AVs) have garnered significant attention. Principle work has centered on refining the technical possibilities of AVs (Cordts et al. 2016; Castrejón et al. 2017), scoping the legal restrictions on AVs (Science and Technology Select Committee 2017; Law Commission 2018), and negotiating the utilitarian ethics of AVs (Bonnefon, Shariﬀ, and Rahwan 2016; Noothigattu et al. 2017). However, less time has been dedicated to the politics of navigational decision-making integral to AVs: how navigational decisions are made, what spatial knowledges are employed, and what calculations these vehicles are dependent on to make such decisions (Sprenger 2018). In this paper, I argue that three navigational practices are integral to autonomous driving worlds. I refer to these as route-calculation, terrain-optimization, and object-recognition.

As Reyner Banham suggested, ‘the car-borne view is neither detached nor in parallax. The observer plunges continuously ahead into a perspective that is potentially dangerous and demands his active attention’ (Banham 1972, 243 as quoted in Merriman 2007, 14). Calculation of these vital moments in advance of the present is a critical capacity of the human driver. By extension any AV will be required to capture, compute, and respond to, a ‘potentially dangerous’ perspective wrought by being in control of a vehicle. As Banham suggests, this demands ongoing ‘active attention’. I suggest that there is a pressing political, ethical, and socio-technical need to address the automobile ‘space-times of decision-making’ (McCormack and Schwanen 2011).

First, as part of the ‘driver–car’ (Dant 2004) relationship, the ‘car-driver hybrid’ (Sheller and Urry 2000), ‘intelligent transport systems’ (Urry 2004), or ‘intelligent traffic’ (Beckmann 2004), I suggest that maps have always been integral to driving a car. However, with the quasi-forms of the car-driver or
driver–car dissolving into wider, distributed automobile systems, in which humans are downgraded to supervisors rather than operators, I introduce the term ‘driving-machine’.

Comprising of many disparate parts and actors, in which anticipatory forms of navigational decision-making become critical (Hind 2016), the term connotes the generative, dynamic, preemptive and reactive, multi-agential and infrastructural form of AVs that mark it as distinct from a simple car-driver–driver–car hybrid. Here I take the driving-machine as a moving, geographical ‘site’ (Hanson 2018), in order ‘to trace the extensiveness of [the] object’ and to ‘attend to the spatial complexities of automation’ (Bissell 2018, 60).

In unpacking the driving-machine, I suggest that research into the automation of industrial processes as well as aircraft fly-by-wire systems identifies a fundamental paradox, as well as the many ‘ironies’ (Bainbridge 1983) of automating manual processes such as driving a car. As tasks are removed from a human operator and reassigned to automated systems, risks and responsibilities shift rather than disappear. I argue that acquired driving skills such as navigation will not be lost, but rather reassigned, recalibrated, or even intensified.

I proceed by considering each of these navigational practices in turn. Route-calculation is the first practice that has already been significantly automated. Mobile apps such as Waze have engendered novel social interactions between drivers, and throughout the driving world (Hind and Gekker 2014). AVs will shift the attentional form of driving and navigation, valuing feedback loops of data generated by the act of being autonomously driven, precipitating the rise in navigational supervision.

Terrain-optimization is the second navigational practice that I argue assumes importance in any AV world. Ideal environments, or terrains, have long-shaped the design and production of automobiles, as well as the navigational possibilities of vehicles. However rather than terrains being shaped or ‘optimized’ for AVs, I argue that AVs are being computationally optimized for terrains. Here, terrain is more than terra firma, incorporating the sensing of elemental (McCormack 2016), atmospheric (Anderson 2009; Durham Peters 2015; Hansen 2015), and meteorological phenomena. Terrain-optimization shapes the AVs ability to pay attention. This optimization means that AVs are suited to operating in road environments, such as on motorways, where specific driving tasks (such as lane-changing) can be successfully performed.

Object-recognition is the third navigational practice I wish to discuss. While historically a task for human drivers, AVs are beginning to possess object-recognition capabilities in which ‘active phenomena’ (Hind 2016, 207), such as pedestrians, animals, and other vehicles, can be sensed. This object-recognition process has generally consisted of a discretization of things-in-the-world into objects, as well as a categorization of such objects for navigational decision-making purposes. Although this two-step process is being tested in a variety of ways, with an assortment of sensor technologies and parameters, there are common technical and ethical questions being asked concerning these regimes of valuation (Gerlitz 2016).

I draw on a wide variety of cases: from historical autopilot technologies and car advertisements to robotics competitions, ongoing AV research projects, and speculative technological proposals, to support the argument that AVs are transforming automobile navigation. I conclude by suggesting that paying attention to how these driving-machines shape space, and affect navigational decision-making, is a critical part of understanding the effect of AVs on society more generally.

Navigation, automation, and supervision

In 2013, the American journalist Adam Fisher wrote an article in the New York Times Magazine called ‘Google’s Road Map to Domination’. In it he struck an anxious tone, suggesting that:

> We’re fast approaching an endgame in which the capacity to read a map could become a lost art. The online-map era started with a flowering: Rademacher’s HousingMaps.com. Foursquare and others took the concept to its logical conclusion. It’s no exaggeration to describe the smartphone as the equivalent of a cursor moving through a one-to-one-scale map of the world. Today, turn-by-turn navigation is the quintessential map app.
In a potted history of Google’s mapping exploits – from early mash-ups to autonomous driving – Fisher supposedly identified a trend toward the absorption of the map ‘into the machine’. Map reading, he argues, ‘could become a lost art’ as navigational skills evaporate with the automatic production of space (Thrift and French 2002).

There are two points to make about Fisher’s proclamation. First, that mapping – and more precisely, vehicular navigation – has always been a hybrid endeavor. The map has always been part of the machine, from early photo-auto guides (Thielmann 2016) through A-to-Z road atlases (Wood 2010), to satellite navigation systems or ‘satnavs’ (Brown and Laurier 2012), and social navigation apps (Hind and Gekker 2014). Hybridized, they become part of the ‘driver–car’ (Dant 2004) or ‘car-driver’ (Sheller and Urry 2000; Beckmann 2004). These terms acknowledge the human labor involved in crafting, maintaining, managing, and operating such machines; positing the vehicle as an ‘assembled social being’ (Dant 2004, 74) and as a ‘quasi-object’ (Latour 1993) in which human subjects (drivers, passengers, pedestrians, etc.) are not ‘conceived as autonomous from … all conquering machinic complexes’ (Sheller and Urry 2000, 739). Navigation, in this sense then, is not a human skill threatened with extinction, but as a hybrid capacity always-already constitutive of the driving-machine.

Secondly, that full automation – when ‘cars drives themselves’ – is improbable. Evidence from the automation of industrial machinery and from aircraft fly-by-wire systems identifies an ‘automation paradox’ (Bainbridge 1983), in which greater automation begets greater human oversight. This in turn results in an even greater importance on human skill and intelligence as the possibility of concatenated, complex, and catastrophic errors increases.

Concerning the automation of industrial processes, Lisanne Bainbridge (1983, 777) argues that there are a number of ‘ironies’ to automation. One of these is that the more technically advanced and autonomous a control system is, the more skilled the system monitor must be. As a ‘formerly experienced operator’ (775) is turned into a machine monitor overseeing rather than generating industrial activity, cognitive skills and manual expertise are lost. Work processes then become reliant on residual knowledge accrued by former manual operators. Such knowledge is not, in Bainbridge’s account, similarly gained by new human monitors. Yet, ‘[b]y taking away the easy parts of … [a] task, automation can make the difficult parts of the human operator’s task more difficult’ (777), as they are forced to respond to machinic failures without working knowledge of the system. Any drive toward automation results in not only a recalibration but an intensification of human skill-acquisition. Nissan’s proposal for human ‘mobility managers’ to decide on behalf of AVs when they encounter unexpected situations is one such example (Nissan 2017) of this supervision in action.

Aircraft fly-by-wire systems also present some interesting precedents. Fly-by-wire is an electronic system that mediates certain flight actions. Most notably, it allows the flying of aircraft according to specific performance envelopes, limiting the range of control movements that can be made (aileron positions, angle of attack, etc.). Fly-by-wire systems dovetail with advanced dashboard instrumentation, such as flight management systems, for pilots to manage control features safely (Mingle 2015). However, by doing so, manual operators (pilots) must be trained in recognizing instrument signals. When a fly-by-wire system enters a different mode – such as during the fatal Air France Flight 447 crash in 2009 – pilots must interact with it differently. Yet ‘mode confusion’ (Mingle 2015, n.p.) is a common occurrence, with pilots failing to understand an aircraft’s actions. AVs, as accidents during testing suggest (Levin and Carrie Wong 2018), are equally demanding of these novel supervisory skills.

Driving and navigation have an entwined relationship, with map reading unlikely to become a lost art. On the contrary, it is likely that the automation of driving will lead to an intensification of forms of navigational supervision – recalibrating and redistributing navigational practices. Evidence
from the automation of industrial processes and the introduction of fly-by-wire technology poses an ‘automation paradox’ (Bainbridge 1983) to automobility, despite the hopeful proclamations. In the next section, I detail the first of three navigational practices I argue are integral to autonomous driving worlds: route-calculation; terrain-optimization; and object-recognition.

**Route-calculation**

Route-calculation is the most common of the practices I wish to discuss. It refers to the way road and navigational data are utilized to generate recommended ‘A-to-B’ journeys for vehicle users, allowing them to follow automated instructions turn-by-turn. In some sense, vehicle route-calculation has been automated since the early twentieth century, with the invention of ‘photo-auto guides’ (Thielmann 2016). These handy publications, produced by Rand McNally and other road map publishers, comprised of photographic snapshots of intended journeys, annotated with directional arrows, accompanied by textual descriptions. These were no doubt turn-by-turn navigational devices, designed to aid the driver in reaching their destination ‘without a hitch’ (Latour 2013, 77).

In 1999, the British Automobile Association (AA) launched its online AA Route Planner, constituting the first digital automation of route-calculation. Prior to this date, it ‘mailed an average of 250,000 routes to members each year’ (AA 2012, n.p.), composed by AA employees according to specific navigational requests by its members. Satnavs are now commonly used by drivers to provide such turn-by-turn navigation, since global positioning system (GPS) was made commercially available in 2000. With these systems, drivers have been able to input destinations into a mobile, dashboard-mounted device, and drive according to visual and audio commands. However, since the late 2000s, mobile apps such as Waze have shifted the satnav experience significantly, contributing both to the digital socialization of driving and to the further automation of route-calculation. I discuss each of these in turn with reference to Waze.

First, ‘social navigation’ apps (Hind and Gekker 2014) such as Waze have populated an empty satnav world with other road users. While driving has always been a social activity (Featherstone 2004; Sheller 2004; Laurier 2004; Brown 2017), using a traditional satnav is limited, in this respect, to interaction with the satnav assistant and fellow possible passengers. This constitutes what Barry Brown and Eric Laurier (2012, 1) refer to as the ‘normal, natural trouble’ of driving with a satnav system. Waze engenders a fundamental shift in the navigational experience: adding the user-driver into a digital world populated by a specific type of road user: the ‘Wazer’. It results in a new phase in digital navigation – in which the digital is made social – but also constitutes a new stage in the digitalization of both the social space of the car and the social space of the road (Brown 2017).

Second, like other digital platforms, the data produced through interaction with the platform are valued and utilized. In Waze, movement data are integrated into route-calculation by monitoring user activity in, and through, the app. GPS, for example, must be activated for Waze to work. Thus, as Alex Gekker and I have suggested:

> Waze users contribute – knowingly and unknowingly – through active driving, desktop editing and passive metadata collection … The data gleaned helps to not only build up a vast picture of the journeys made with Waze, but also the state of the road network in general. (Hind and Gekker 2014, 7)

The circulation of data collected ‘knowingly and unknowingly’ from users is arguably integral to the AV’s ability to gauge current, and future, road conditions to calculate optimal routes.

The relationship between ‘active’ and ‘passive’ in an autonomous driving world complicates Banham’s idea of active attention, qualifying the attentional shifts identified by automation and safety researchers in other domains (Bainbridge 1983; Stanton and Marsden 1996; Janssen, Wierda, and van der Horst 1995). In an autonomous driving world, the need for active human attention is likely to shift toward more passive, or intermittent attentional requirements, as seen on factory floors or in aircraft cockpits.
The need for active, continued, and consistent attention is likely to become the responsibility of the vehicle and its integrated sensor, navigational and computational systems. In this scenario the human driver, like Bainbridge’s (1983) machine operator, now assumes the role of (mere) supervisor. The ‘real-time’ calculation of the driving world, as rendered by vehicle manufacturers, would arguably be made possible by passive data collection generated through nonhuman driving (Stilgoe 2018). Relieved of immediate driving duties, and with it the need for particular attentional capacities, vehicle users in an autonomous driving world are more likely to engage in the kinds of tasks Wazers do now: adding data such as accident locations onto the map, communicating with other drivers, and providing general diagnostic route assistance. Finally freed from the need to pay active attention, human drivers will arguably settle for becoming route-supervisors while route-calculation is administered primarily by the AV.

The calculation of driving routes has been at least partly automated since the early twentieth century, with the digitization of route planning taking place in the late 1990s. More recently, social navigation apps such as Waze have digitized social relations, allowing road users to communicate with each other through their mobile devices. The crowdsourcing of road data – everything from accident locations to incorrect speed limits – has radically reorganized where and how people drive (Foderaro 2017; Lopez 2018). This reorganization has been built on the generation of massive amounts of data derived from active human driving and passive data collection. Automation threatens to relieve human drivers of active driving and thus the requirement to remain actively attentive. Instead, they are likely to become route-supervisors, overseeing – rather than administering – the turn-by-turn navigation of an AV. These antecedent developments are vital to understanding the contemporary development of the navigational capacities of AVs.

**Terrain-optimization**

Terrain is a familiar term in the automobile industry, and a variety of factors affect how vehicles cope with different terrains. Legislative restrictions regulate absolute vehicle sizes (Department for Transport 2017), market segmentations distinguish between potential buyers (executive, family, leisure), and aesthetic (body shape, body material) and technological (engine, drivetrain) features determine the performance limits of a vehicle. Four-wheel drives, All-Terrain Vehicles, and Sports Utility Vehicles, for example, clearly denote or subtly imply the terrain suitability of such vehicles.

AVs are also dependent on what I refer to as terrain-optimization. However, this is not the optimization of terrains for vehicles, such as the ‘designing in’ of material infrastructures to deal with AVs (Stilgoe 2017). Instead, terrain-optimization refers to the optimization of vehicles for terrains. This ontological reversal affords the AV agency in the cultivation of new kinds of terrain, possessing the ability to bring new driving worlds into being, shaping the conditions of their use. Optimization, in this sense, is when algorithms are used to make a computational process more efficient. Dijkstra’s shortest-route algorithm is one such navigational example (Lanning, Harrell, and Wang 2014). This optimization, rather than a kind of base map, is more an active mapping of elemental phenomena (McCormack 2016) that simultaneously enables the recognition of objects.

In this section, I argue that each driving-machine renders terrain differently, based on their sensory capacities. While some are optimized for motorways, others are tailored to urban and suburban environments. A complex arrangement of sensor technologies in each machine – everything from cameras to radar, LIDAR, and GPS – results in a complex sensing of road surfaces, lines, signs, and meteorological conditions. As Sebastian Thrun, Mike Montemerlo, and Andrei Aron (2006, 1) suggest; ‘[t]he ability to perceive and analyze terrain is a key problem in mobile robot navigation’, even if many driving tasks are now considered ‘solvable problems’ (Stilgoe 2018, 31) by AV developers. I end this section with an example of terrain-optimization in action.
**Terrain not territory (nor landscape)**

Terrain is not the same as territory. As Elden (2010, 809) has argued, territory inscribes power, historically involving a ‘calculative grasp of the material world’ with the help of cartographic techniques and technologies such as triangulation and theodolites. Maps and territory therefore have a shared, entwined history, in which sensing, surveying, mapping, and navigating inscribe and manifest territory (Law 1984; Edney 1997; Akerman 2009). Yet for Elden, territory itself can be considered a ‘political technology’ (2010, 810), in which strategic and technical work is undertaken to abstract, secure, manage, and therefore shape space and spatial relations. Similarly, work by Del Casino and Hanna (2006) and Kitchin and Dodge (2007) has sought to dismantle an ontological divide that has persisted between the map and the territory, the technology and the world. As Hind and Gekker (forthcoming, 13) have proposed, AVs are likely to ‘further deepen the collapse of the map into the territory’, resulting not in another example of a ‘co-constitutive’ relationship between map and territory, but an entirely ‘flat ontology of vehicular navigation’ (1), in which the map and territory fuse as one. Territory is created when geographical phenomena that comprise terrain are captured and mapped (Wilson and Elwood 2014).

Terrain is also not a synonym for landscape. Peter Merriman (2007, 12) argues that ‘[d]riving is not solely a visual experience’, with many sensory aspects critical to the manual operation of a vehicle, as others have noted (Dawson 2017; Pink, Fors, and Glöss 2017). Yet landscape connotes a romantic vision of the driving experience; reserved only for ‘back-seat passengers’ and at odds with the ‘plunging perspective’ (14) of the driver and the calculative vision of the AV’s assorted, assembled, and integrated sensing technologies. As some have intimated (Rose 2018), these technologies do not ‘see’, and therefore do not generate the kinds of perspectives associated with driving through landscapes. Moreover, when human operators become navigational supervisors of the driving-machine they arguably fuse the passive, parallax view of the back-seat passenger with the previously active perspective of the driver. In any case, landscape – as a kind of holistic, aesthetic, cultural form of visioning (Cosgrove 2008) – is not enrolled into any optimization process.

Terrain is also more than terra firma. It is not just the hard ground or solid earth, but a voluminous incorporation of all things elemental (McCormack 2016) or atmospheric (Hansen 2015). It is, as Elden (2017, 199) contends ‘where the geopolitical and the geophysical meet’. In other words, ‘[t]errain can be land, water or some blurring of the two in indeterminate and dynamic environments …’ (201–2). As the driving-machine takes control, it must actively capture, sense, store, and calculate the properties – and capacities – of the driven terrain, everything from road material, width and markings to meteorological conditions such as rain, fog, and sunshine. These meteorological conditions interact with the ‘hard’ ground or ‘solid’ earth to fundamentally reconstitute them. As a Bloomberg article recently suggested, ‘self-driving cars can handle neither rain nor sleet nor snow’ (Stock 2018, n.p.), due to the effect each condition has on the road surface and the AV’s sensory capacities. ‘Cameras’ the article continues, ‘are useless in fog and heavy snow’, while ‘lidar lasers career wildly off raindrops and snowflakes’ (n.p.). A fatal crash involving a semiautonomous Tesla vehicle was also due to its inability to ‘notice … the white side of [a] tractor trailer against a brightly lit sky’ (Tesla 2016, n.p.). AVs enroll these elemental properties into decision-making, with the road environment being constantly sensed and acted-upon. As such, I argue that the sensing of the elemental properties and capacities of terrain are critical to the novel decision-making power of the driving-machine.

**Optimization**

Michael Dieter argues that the genealogy of optimization can be traced back through the ‘formalization of decision-making in the context of operations research (OR) during the Second World War’ (Dieter 2017, 72). This, he continues, ‘arose as a kind of auxiliary apparatus of management, wherein computational procedures and protocols mediate settings of institutional
judgement’ (72). Here, the AV not only seeks to calculate an optimal route (fastest, shortest) likely using a variant of Dijkstra’s algorithm (Lanning, Harrell, and Wang 2014; Byrne 2015) but also seeks an optimal driving state (smooth, resourceful). This optimization process, therefore, is dependent upon the optimal sensing of terrain, without which neither optimal route-calculation nor an optimal driving state can be achieved. This algorithmic process is dependent upon the capture, recognition, classification, and prioritization of sensed data (visual, air pressure, meteorological) derived from various sources (roof-mounted LIDAR, tire sensors, windscreen sensors).

Most vehicles, autonomous or otherwise, can be driven in a wide variety of environments. However, only with the development of autonomous features are they actively required to refine and react to the navigation of terrains (Mitchell 1986). The gestural and social dynamics that comprise manual driving need to be codified in respect to these spaces (Brown 2017). Expected – and unexpected – driving environments will present the vehicle with notable challenges it must cope with. These challenges may equally comprise of ill-defined road markings, or foggy conditions.

The optimization of AVs affects the vehicle’s ability actively to pay attention. This optimization process can be carried out in numerous ways. Thrun, Montemerlo, and Aron (2006, 1) suggested the use of a ‘Probabilistic Terrain Analysis’ (PTA) algorithm for ‘terrain classification’. It is this classification that doubles as optimization, with the consistent tweaking, refining, and adaptation of classifications resulting in ongoing optimization. In their words, ‘[t]he PTA algorithm uses probabilistic techniques to integrate range measurements over time, and relies on efficient statistical tests for distinguishing drivable from non-drivable terrains’. PTA is concerned with adjusting errors recurrent in the sensing process, to determine the navigability of terrain.

In this case, the PTA algorithm was used in the 2005 DARPA Grand Challenge, an annual robotics competition hosted by the US Defense Advanced Research Projects Agency (DARPA), assisting the winning AV in negotiating a 132-mile course. The algorithm, as Thrun, Montemerlo, and Aron continue, ‘processes range data acquired by the single-axis laser scanner mounted horizontally on a moving robotic platform’ (1). These ‘range data’ – the primary mode through which terrain is captured or sensed – are then used to generate a ‘2D environment suitable for robotic driving’ (1). Through a ‘rule learning’ process (Stilgoe 2018, 29) it uses its own data to refine and optimize the vehicle’s working parameters. As Ethan Alpaydin (2016, 149) has more recently suggested, ‘[m]achine learning plays a significant role in self-driving cars’, with vehicles using a constant stream of terrain data to make decisions.

HERE, a mapping company now owned by a consortium of German vehicle manufacturers, also recognizes the need for AVs to understand terrain as voluminous (Elden 2013) – and, by extension, elemental or atmospheric. Their Live Roads technology, for example, is designed to feed terrain data from the vehicle into AV decision-making:

Let’s say, for example, that in a particular area tire sensors on certain cars report that the tires are slipping. Meanwhile other cars close by send information that the windshield wipers are on. At the same time the local weather agency sends out an alert that temperatures have dipped below freezing. Live Roads aggregates and analyzes all of that information to understand that there is black ice in a certain area and can then send that information back to all cars headed there. (Rayasam 2015, n.p.)

HERE’s Live Roads system is designed to enroll atmospheric measurements – from tire sensors, windscreen wiper sensors, and metrological agency alerts – into the calculative capacities of the AV. This combines the typical Dijkstra’s shortest-route algorithms used in navigational software, with a dynamic, algorithmic optimization of terrain. In turn, this data should theoretically be able to affect how the car moves, the speed at which it is traveling, the status of other, optional driving features, or even the entire route taken. As HERE suggests, this isn’t simply the sensing of grounded objects – the ‘rocks, vegetation, berms, ruts, cliffs, overhangs … and man-made artifacts …’ identified by Thrun, Montemerlo, and Aron (2006, 1) – but
varying other atmospheric agents. Maintaining a vigilance toward these phenomena, therefore, is a critical component of terrain-optimization.

I will now present an example terrain currently being optimized and rendered differently by semi-AVs: the motorway.

**Tesla and the motorway**

Tesla is an American car manufacturer founded in 2003. It currently makes several all-electric vehicles, predominantly for the luxury market. In the second quarter (Q2) of 2018, it manufactured 53,339 cars (Tesla 2018a, n.p.). At the end of Q2, there were 420,000 net reservations for its Model 3 vehicle (n.p.), which has been in production since 2017. In 2014, it launched a driver-assist feature called Autopilot, which used an array of sensor systems (cameras, radar, etc.) to enable adaptive cruise control, lane departure warnings, and the partial automation of other driving tasks. Since then, its ‘Enhanced Autopilot’ feature enables a Tesla to perform an increasing array of functions, including to:

- Match speed to traffic conditions, keep within a lane, automatically change lanes without requiring driver input, transition from one freeway to another, exit the freeway when your destination is near, self-park when near a parking spot and be summoned to and from your garage. (Tesla 2018b, n.p.)

While the various models Tesla manufactures can be driven in many terrains, under different conditions, Tesla has developed its Autopilot system to attend to problems encountered during longer-distance, motorway driving. As noted in the promotional material above, Tesla’s Enhanced Autopilot deals with a select array of driving tasks and problems typically encountered on a motorway. Of the seven listed above, five involve maneuvers on the motorway. Further that:

- Once on the freeway, your Tesla will determine which lane you need to be in and when. In addition to ensuring you reach your intended exit, Autopilot will watch for opportunities to move to a faster lane when you’re caught behind slower traffic. When you reach your exit, your Tesla will depart the freeway, slow down and transition control back to you. (Tesla 2018b, n.p.)

This is not to say that Autopilot does not work in other environments, in different terrains. However, it is to say that the software itself is being designed, developed, tested, and optimized with motorways in mind. More attention, therefore, has been placed on resolving attentional and navigational issues on the motorway than in other environments.

The motorway is the historic home turf of auto-piloting technologies. Chrysler’s ‘Auto-Pilot’, developed on a test run between Detroit and New York City in 1958, was one of the first automated driving features. In one advert, it was touted as ‘one of the greatest automotive inventions ever developed’, and otherwise as an ‘amazing new device’, ‘acclaimed by experts’ (Chrysler [1958] 2018, n.p.). It was described as having three benefits, able to: automatically warn the driver of a dialed speed limit; maintain the speed of the vehicle without the need to keep a foot on the accelerator; and save as much as ‘three gallons out of every tankful!‘ (n.p.). As suggested in the same advert, if the buyer did ‘considerable driving’, then Auto-Pilot would ‘more than pay for itself in gas economy alone’ (n.p.). While there was no technical, nor legal, restriction on using this system on other roads, Chrysler’s Auto-Pilot was clearly developed with motorways in mind.

The apotheosis of Tesla’s Autopilot, following this logic, should be the successful navigation of a ‘stack interchange’. As Peter Merrington and Ilana Mitchell (2017, n.p.) have noted, the stack interchange ‘enable[s] traffic to change direction from any point without deceleration’. Their design minimizes (although does not eradicate) ‘weaving’, in which drivers move quickly across motorway lanes to exit carriageways (De Blasiis et al. 2016; Kusuma et al. 2014). The maneuver requires drivers to pay careful attention to avoid other vehicles. Stack interchanges, therefore, ensure a smooth traffic flow in which average motorway speeds are maintained.
Thus, the successful autonomous piloting of a stack interchange is a plausible, surmountable challenge for Auto-Pilot, considering it is designed to execute lane-changes, transitions, and exits. Equipped with this navigational ability, it is arguably designed specifically to negotiate these kinds of terrains.

**Object-recognition**

AVs need to perceive the social world around them. As Levinson et al. (2011, 3) suggest, ‘[i]n some scenarios, an AV requires deeper understanding of the environment to behave correctly …’. Yet, unlike other navigational practices, object-recognition has historically been the responsibility, and an acquired skill, of the human driver. As John Urry (2004, 31) suggests, human drivers are primed in various ways to respond to emergent, possible threats: ‘[e]yes have to be constantly on the look-out for danger, hands and feet … ready for the next manoeuvre … body gripped into a fixed position’.

Object-recognition, then, is a sensory orientation toward the near driving future; an eye on the horizon of possibility, an awareness of the world around and beyond. Machinic object-recognition capabilities, equally, require AVs to sense and ideally respond to a full gamut of other active, dynamic agents and objects on, near, around, and beside the immediate road environment (Jain 2004). These may include pedestrians, cyclists, animals, other vehicles, or temporary construction features such as traffic cones. The automation of this heretofore composite, multisensory human driving skill is no easy task, requiring the generation and analysis of large volumes of data. Various estimates suggest AVs will use 4000 GB of data per day (Krzanich 2016) and upwards of 300 TB per year (Dmitriev 2017). Such data, like that generated through terrain-optimization, are also used in AV machine learning.

As I have argued, terrain-optimization is not a base map, but an active mapping of elemental phenomena that enables object-recognition. Without the optimization of the driving terrain, objects cannot be located and, as such, object-recognition cannot be performed. It is the former practice (terrain-optimization) that simultaneously sets the parameters for a live response to other road users and things (object-recognition). It is an ongoing precondition for object-recognition dependent upon the sensing of traversable terrain, the rendering of a viable driving world, in communication with a desired, calculated driving route.

Driving entails the negotiation of a social world replete with possible risks. As John Urry (2004, 29) says:

> Junctions, roundabouts, and ramps present moments of carefully scripted inter-car-action during which non-car users of the road constitute obstacles to the hybrid car-drivers intent on returning to their normal cruising speed, deemed necessary in order to complete the day’s complex tasks in time.

Banham’s notion of active attention is drawn into focus again. In the UK, human drivers must pass a ‘hazard-perception’ test (Driver and Vehicle Standards Agency 2015) in order to demonstrate their ability to assess, evaluate, and respond to emergent risks. While some of these risks are embedded within the terrain (speed limits, traffic lights), others are contingent upon things passing through it (vehicles, pedestrians). For an AV, each encounterable object constitutes a kind of latent hazard. Every object poses a potential risk, with this potentiality made formalizable by the driving-machine.²

**Discretization and categorization**

This recognition process entails two steps. First, it requires that things in the world are considered as discrete objects. While this casts doubts on the ‘discrete-ability’ of the world, it also raises questions of valuation and worth (Boltanski and Thevenot 1991; Stark 2009; Adkins and Lury 2012). As Carolin Gerlitz has suggested, social media data are ‘multivalent’, capable of juggling multiple
'valuation regimes' (Gerlitz 2016, 23). Similarly, David Stark (2009, 108–109) has argued that certain kinds of economic activity cultivate ‘bountiful friction’ between different forms and modes of valuation. However, I argue that multiple value regimes need to be smoothened when recognizing objects. In other words, value conflicts need to be resolved, or at least prioritized when recognizing objects. Any ‘bountiful friction’ is likely to generate an interpretive fiction of the driving world, and with it, possibly disastrous consequences.

Second, it requires that things in the world can be categorized into particular object classes, defined by their form and characteristics. The semantic labeling of street imagery is one approach to categorization. In this, large visual databases are compiled with ‘a diverse set of stereo video sequences recorded in street scenes from 50 different cities’ (Cityscapes 2017a, n.p.). Through such a database, the world can be divided into general groups and specific classes. In the Cityscapes dataset, groups include human, vehicle, and construction. Classes include specific human types (person, rider), vehicle types (car, truck, motorcycle), and construction types (building, wall, fence). Each class is defined to aid interpretation.

What is notable is how objects may switch classes over time or be subdivided into separate classes. Dynamic objects, for instance, might easily transform into static objects (a dead animal), or maintain stasis for an indefinite period (fly-tipped sofa). Traffic signs, for example, are divided according to which side the information is visible on, with the other plain side categorized separately. Commercial signs attached to buildings are technically not signs at all, but buildings. These categorizations matter because they affect how an AV’s sensory system reacts to the social world around it.

Categorization by group and class may be performed for two reasons. First, to assess ‘the performance of vision algorithms for two major tasks of semantic urban scene understanding: pixel-level and instance-level semantic labelling’ (Cityscapes 2017b, n.p.). Second, for ‘supporting research that aims to exploit large volumes of (weakly) annotated data, e.g. for training deep neural networks’ (n.p.). Put otherwise, it is to enable algorithmic object-recognition in the first place, and to refine the recognition process accordingly, by way of improving the robustness of machine learning. Semantic labeling, and the subsequent discretization and categorization of objects, is necessary for driving-machines to make sense of the world. Without this they would make arbitrary navigational decisions based on no identifiable criteria.

Object-recognition may rely on pixel labeling, the identification of objects according to the pixels it occupies on a screen. Here, the arrangement, size, color, and density of pixels are used to determine an object (Fridman et al. 2017; Huang et al. 2018). Other semiautomated approaches identify objects by polygon annotation (Castrejón et al. 2017). This approach works by drawing object outlines, the peculiar shape of a cyclist or the rectangular form of a bus. According to Castrejón et al. (1), this latter process has sped up annotation by a ‘factor of 4.7 across all classes’. These two parallel approaches – one focused on pixel recognition and the other on polygon outline – mirror the divide between raster-based image processing and vector-based recognition. Efforts by Castrejón et al. continue long-running efforts to automate the image classification process in Geographical Information Science, remote sensing, and other cognate disciplines (Lu and Weng 2007). As such, object-recognition has long been a digital cartographic concern.

**Errors**

As Ethem Alpaydin (2016) suggests, there are two types of errors pertinent to algorithmic classification: false positives and false negatives. In his example:

If the system predicts cancer but in fact the patient does not have it, this is a *false positive* – the system chooses the positive class wrongly. This is bad because it will cause unnecessary treatment, which is both costly and also inconvenient for the patient. If the system predicts no disease when in fact the patient has it, this is a *false negative*. (53, authors’ emphasis)
Thus, returning to the DARPA Grand Challenge, algorithmic false positives constitute the identification of objects not actually present. This is ‘bad’ because the vehicle may react inappropriately, and drive into genuinely dangerous terrain. False negatives are similarly bad because the vehicle may not sense actual, present objects – further leading it into a dangerous situation.

Returning to Thrun, Montemerlo, and Aron, what made Stanley so successful was the PTA algorithm’s ability to ‘distinguish between actual obstacles and “phantom” obstacles’ (Thrun, Montemerlo, and Aron 2006, 1). It is these ‘phantom’ obstacles – a result of sensor errors – that pose a considerable problem for AVs. Yet:

In one dataset, it [the PTA algorithm] reduce[d] false-positive error rate from 12.6% to 0.002% without significantly affecting the false-negative rate. Such numbers mattered greatly for the DARPA Grand Challenge: false-positives correspond to “phantom obstacles” that . . . easily mislead the robot into hazardous terrain. (1)

Unlike, Alpaydin’s medical example, it might appear that the outcomes here for both false positives and false negatives are roughly the same: the AV is lead into a hazardous environment. Nevertheless, what mattered in the DARPA Grand Challenge is markedly different from what matters in real-world driving situations. Objects in the wider social world are not obstacles in a race. Phantom obstacles in the DARPA Grand Challenge fit the features of an arid terrain: rocks, shrubs, depressions, landforms. These would typically fit into a single group in the Cityscapes dataset (Nature), rather than be spread across multiple groups and classes. In a social world, decreasing the false positive error rate likely increases the possibility another such object – a person, an animal, another vehicle – is hit, by restricting the criteria, thus increasing the threshold, for a ‘real’ object.

Elaine Herzberg was killed by an AV in Tempe, Arizona because she did not reach this threshold (Gibbs 2018). While the vehicle is alleged to have nominally detected Herzberg, it mis-categorized her as an ‘object’ unworthy of evasive action. Thus, we see that in a social world, rather than in a race such as the DARPA Grand Challenge, that false positives matter. As David Bissell (2018), James Ash (2018), and Jack Stilgoe (2018) have all noted, it is critical to recognize the politics of this sensing. These errors have social consequences.

The calculation of the likelihood of (near) future outcomes, plus their assigned possible value, thus comprises ‘expected value’ calculations (Alpaydin 2016, 54), in which the values of each outcome are judged against other possible outcomes. The question for autonomous driving is whether decisions on optimization and classification can be rejected or deferred to a human; who or what has the ‘imperative to decide’ (McCormack and Schwanen 2011, 2809)? Moreover, can these critical navigational decisions be executed decisively, live and on the move?

I will now present one case that demonstrates the value judgements implicit in developing object-recognition capabilities.

Waymo and the (sub)urban

Waymo, Google’s AV project, has built a piece of software known as Carcraft (Madrigal 2017, n.p.). It is in this software that various real-world and simulated tests are carried out. A related software called XView offers Waymo employees the opportunity to see ‘what the car is “seeing”’ (n.p.). This live application shows object-recognition in action, depicting an array of discrete things as ‘little wireframe shapes’, colored according to their categorization in different object groups (n.p.). Object-recognition enables this world to be navigated and driven. Waymo has generated 20,000 scenarios from its structural testing, involving such object-recognition (n.p.).

Alongside this Waymo has also completed nearly 636,000 miles of testing in California throughout so-called ‘more complex urban and suburban environments’ (Korosec 2017, n.p.). This is a marked difference from the approach taken by Tesla. Motorways do not constitute the primary test environment for Waymo, so-called ‘complex’ urban and suburban environment do.
Unlike on the motorway, (sub)urban environments exhibit a rather more unruly set of social phenomena outside the vehicle. Even more specifically, Waymo has been conducting tests in complex Californian urban and suburban environments. As well as on public roads, tests have taken place in a closed Californian facility, with ‘real’ junctions built – for real – at the facility itself, so its AVs can be driven through these junctions, and a multitude of scenarios played out. Many of these junctions, as Madrigal (2017) attests, have been built because Waymo employees have encountered them in real life.

Here, we are faced with two questions. First, what is valued? Madrigal notes that Waymo uses a ‘prop stash’ replete with inanimate objects: ‘dummies, cones, fake plants, kids’ toys, skateboards, tricycles, dolls, balls, doodads’ (2017, n.p.). These objects are used in scenarios to test the object-recognition capabilities of the AV. If facial-recognition technologies and their attendant algorithmic functions are known to discriminate and consolidate institutional bias (see Mittelstadt et al. 2016; Ananny 2016; Amoore and Piotukh 2016), then what effect does the prop stash have on semantic labeling, image classification, and machine learning overall? What has the right to enter the sacred prop stash? What deserves to be classifiable and therefore recognizable? Each of the objects listed by Madrigal has their own specific properties and capacities. Each fake plant, kids’ toy, and tricycle can be categorized, prioritized, and valued differently depending on its form and potential, as well as the method employed (pixel labeling, polygon outlining) to classify it.

Second, and just as critically; what is mis-valued? What happens when the autonomous system categorizes, but mis-recognizes an object? An AV may mis-classify a pedestrian as a rider (by their positioning, posture), mis-categorize a human as an animal (by their dimension, gait, rhythm, speed), or mistake debris as benign vegetation (a plastic bag caught in a cross-wind). The objects that go into such a prop stash matter because it poses a ‘fisherman’s problem’ (Crampton 2002, 15), in which the catch itself ‘furnishes more information about the meshes of [the] net than about the swarming reality that dwells beneath the surface’ (Olsson 2002, 255). A swarming social reality is to be found far beyond the (sub)urban confines of California. The question is whether the AV can recognize it.

Conclusion

As Jörg Beckmann suggested, the ‘car-driver hybrid’ (Sheller and Urry 2000) and, by extension, Dant’s (2004) ‘driver–car’ assemblage were likely to be challenged at some point in the future. That future is arguably now, as work on the sensory (Dawson 2017), phenomenological (Pink, Fors, and Glöss 2017), and design (Pink, Fors, and Glöss 2018) dimensions of driving attest. In this article, I have argued that with the dynamic, multi-agential, computational, calculative AV it is more appropriate to talk of a driving-machine in navigational control.

This paper has sought to address this collapsing hybridity from a navigational perspective. As Beckmann (2004, 90) further suggested, ‘[a]s other activities inside the hybrid are made possible … driving itself becomes a subordinate, a “shadow activity”’. Reference to a driving-machine registers this subordination, signals a machinic takeover of navigational practices, and begins to help explicate the novel ‘space-times of decision-making’ generated by AVs (McCormack and Schwanen 2011). Here the one-to-one relationship between driver and car is made multiple, distributable, with decisions made throughout the sensory apparatus of the vehicle. Consequently, neither the car (as object) nor the driver (as subject) retains even ‘residual’ features as either quasi-objects or quasi-subjects, fused together. The AV, as Beckmann foresaw, fundamentally reconstitutes the relationship between driver, machine, decision, and execution. This corrective – driving-machine over automotive hybrid – is not a dismissal of agency, but an acknowledgment of a changing automotive landscape.

I have argued in this article that three navigational practices will play substantial roles in an autonomous driving world, each a constituent in a new navigational arrangement comprising of
various mapping and sensing technologies. These transformations should be of continued interest to (auto)mobilities, navigation, media, and automation scholars.

First, I have argued that route-calculation will continue to ensure specific, optimal routes are navigated turn-by-turn. While route-calculation is nearly as old as driving itself (Thielmann 2016), more recent shift in the digital socialization of route-calculation (Hind and Gekker 2014) has transformed the navigational experience. While there will be no ‘endgame in which the capacity to read a map could become a lost art’ (Fisher 2013, n.p.), drivers will likely become navigational supervisors, overseeing the calculation and navigation of programmed routes by an AV. Scholars interested in way-finding, navigational knowledges and the effect of new media on cognition, attention, and sociality should find this practice instructive. ‘Map reading’ is an ever-changing skill.

Second, I have suggested that terrain-optimization will enable AVs constantly to fine-tune, manage, and adapt to changing road environments. This is not the optimization of terrains for vehicles, but the optimization of, and by, vehicles for terrains. This acknowledges the AV’s ability to bring new driving worlds into being, shaped by the computational optimization of terrain. Yet rather than just the hard ground, or solid earth, the optimization of AVs for terrain incorporates the sensing of various elemental, atmospheric, and meteorological data. Scholars working on the socio-technical challenges of capturing such phenomena should find this practice intriguing. Terrain is neither stable nor predictable.

Third, I have argued that object-recognition will be required to ensure AVs can not only navigate terrain, follow pre-defined routes but also sense and avoid dynamic, ‘active phenomena’ (Hind 2016, 207) in the driving environment. This active sensing, I have argued, will require the AV to both discretize a diverse array of objects in the world and categorize these diverse objects, correctly, in order to make navigational decisions. It is the object-recognition process that bears a heavy ethical load. Scholars working on the ethics of computation, algorithms, and sensing should be interested in interrogating this practice further. Automobility will be shaped by these capacities.

As Shannon Mattern (2017, n.p.) suggests, ‘[w]ith the stakes so high, we need to keep asking critical questions about how machines conceptualize and operationalize space’. Driving-machines are but one of the many types of machines capable to conceptualizing, operationalizing, and producing space (Thrift and French 2002; Graham 2005; Kitchin and Dodge 2011). More precisely, Mattern demands we interrogate ‘[h]ow … they render our world measurable, navigable, usable, conservable … ’ (n.p.). At this present juncture, with AVs being enthusiastically discussed, publicly tested, and privately developed, there is an opportunity to attend tentatively, provisionally, and speculatively to these questions.

Notes

1. The first four-stack interchange was built in Los Angeles in 1953 and adorns the front cover of Reyner Banham’s Los Angeles: The Architecture of Four Ecologies (Banham [1971] 2009).
2. Both object-recognition and terrain-optimization are kinds of hazard-perception. I do not use the term hazard-perception to avoid conflating the two.

Acknowledgments

Thanks to Brenden Doody and Debbie Hopkins for organizing the Systems of (auto)mobility II session at the RGS 2017 Annual Conference, in which this paper was first presented. Thanks to attendees of the Locating Media Kolloquium for feedback on an early draft. Thanks to Henrik Örnebring for conversations on data and valuation. Thanks to Peter Merrington and Ilana Mitchell for alerting me to the existence of the stack interchange. Finally, thanks to both reviewers for their constructive feedback. Any errors within are entirely my own.
Disclosure statement

No potential conflict of interest was reported by the author.

References


