3. Machinic Sensemaking in the Streets: More-than-Lidar in Autonomous Vehicles

Sam Hind

Abstract
In recent years, lidar has increasingly been deployed in the testing of prototype autonomous vehicles. Rather than mapping forest cover or urban terrain, however, lidar has been used to map driving environments. This chapter explores the machinic sensemaking capacities of prototype autonomous vehicles, both composite as well as "distributed", with various, interconnected sensing systems and software programmes used for orientation, perception, and decision-making. In this, vehicles draw on sensing technologies with different observational ranges, prioritizing some over others at particular distances. Yet enabling this machinic sensibility involves undervalued, and misunderstood, visual responsibilities assumed by so-called "vehicle operators" during tests. Without this important work, prototype autonomous vehicles risk ignoring, or mis-sensing, other road users – with fatal consequences.

Keywords: sensing, machinic sensibility, recognition, distributed media

Introduction

Short for "light detection and ranging", lidar has historically been used for the aerial mapping of vegetation and for surveying urban environments and heritage sites. By emitting pulses of light that bounce back off surfaces and objects, spectral images called "point clouds" are generated, derived from millions of innocuous lidar pulses. In recent years, however, lidar has increasingly been deployed by car manufacturers and technology companies in the testing of prototype autonomous vehicles. Rather than mapping

Rose, G. (ed.), *Seeing the City Digitally. Processing Urban Space and Time*. Amsterdam: Amsterdam University Press, 2022
DOI 10.5117/9789463727037_CH03

Figure 3.1. A stylised rendering of how lidar "sees", or senses, an urban environment. Courtesy of Velodyne Lidar.

forest cover, or urban terrain, lidar has been used as a principal sensing system to map driving environments, and aid the detection of other road users, signs, and lines.

In this chapter, I suggest that lidar is central to the "sovereign" (Bratton 2015; Gekker and Hind 2019; Pasquale 2017) sensemaking capacities of prototype autonomous vehicles, able to "configure territory and power" (Lovink and Rossiter 2019, 99) in new ways. When taken apart, the sensory capacities of prototype autonomous vehicles are both composite as well as "distributed", courtesy of various interconnected sensing systems and software programs used for three critical operations: orientation, perception and decision-making (McCosker and Wilken 2020). Lidar never acts alone; hence I use the phrase "more-than-lidar" to indicate that lidar is reliant upon an integrated suite of sensing systems.

It is often suggested that autonomous vehicles "see" (Davies 2018; Metz 2018; Stilgoe 2017), yet the way they see the world is manifestly different to other forms of (human and non-human) sight. Whilst greyscale point clouds generated by lidar show the world in a skeletal form, equally common technicolour renderings depict it as a kind of parallel hyperreality. Neither capture the urban environment as rendered in photographs, maps, or stylized illustrations (Figure 3.1). Instead, lidar and its ancillary sensing systems render the urban environment anew, in turn affecting how decisions are made within cities.

To address this newness, the chapter will build on Sun-ha Hong's (2016) concept of "machinic sensibility", to consider how autonomous vehicles "sense" rather than see. In this, I suggest that the autonomous vehicle entails four orders of sensing: from *feeling* the shape, texture and form of phenomena in the urban environment, through the rote *capture* of sense data, to the processual calculation of *meaning* from the processing of such data, before arriving at the execution of *good* or acceptable decisions.

The unceasing flow of information that characterize "distributed media" (Munster and Lovink 2005) is rarely the case with prototype autonomous vehicles. Whilst the distribution of machinic capacities can be seen to generate endless successful relays of integration, offering greater fidelity to the sensed environment, this same distribution equally renders relays of *disintegration*. Here, erroneous classifications and clashing system priorities render sensemaking an unevenly distributed activity.

Distinct and *distant* capacities are operationalized through this distribution. In this, distance – most notably, the distance between vehicle and object(s) – becomes a significant spatial principle through which judgements are made, and decisions executed. Yet, for the distant capacities of worldly phenomena to become useful, sensing units within such a distributed system must be *prioritized*, such that some assume greater significance at specific moments, or in specific situations. In this, the capacities of other road users, road surfaces, or entire junctions or road layouts are mobilized in ways that might otherwise not be, with these priorities encoded into the protocols of onboard software.

It is this uneven distribution of machinic capacities that is reflected in the differentiation of "sociotechnical agency" (Rose 2017, 779) at an operational level. As Gabrys and Pritchard contend, sensing practices "shift attention to formations and processes of *experience* across multiple entities" (2018, n.p., emphasis added). As such, this chapter explores how sensemaking in autonomous vehicles generates a differentiation in the distribution of experience, affecting some in qualitatively different ways to others.

As the chapter proceeds, I consider different aspects of the sensemaking capacities of prototype autonomous vehicles. I begin by focusing on the technical features, and operational limits, of specific lidar products used in developmental autonomous vehicles, considering how different models and their possible configurations affect these capacities. I then move on to consider a crash in Tempe, Arizona in March 2018, involving a prototype autonomous vehicle operated by Uber Advanced Technologies Group (ATG), that killed a woman called Elaine Herzberg. I contend that the crash, and the subsequent investigation, revealed the contingencies of classification, as Herzberg was variously re-classified as different objects (car, bike etc.) but never accurately as a pedestrian, in the moments before the crash.

In the final section, I consider how the nominal "supervisor" of the prototype vehicle at the time of the same crash, Rafaela Vasquez, was committed to performing an array of duties meant to enable or "fine-tune" the eventual sensemaking capacities of the autonomous vehicle. By studying the US National Transportation Safety Board (NTSB) report, I query the significance of her own visual sensibilities, and her repeated glances towards the central console of the vehicle. The central console was where her personal mobile phone was allegedly stored, but also where a tablet computer was similarly placed, on which Vasquez was committed to record system errors and driving infractions made by the vehicle in autonomous mode.

Machinic Sensibility

As Gabrys argues, "usually, some version of a cognizing human is at the centre of work on sensing", with sensing "tied to particular types of human embodiment, engagement, and experience" (2019, 724). Nevertheless, as Gabrys continues to suggest that "sensing practices", as she refers to them, extend beyond the human to an often-complex arrangement of "sensing entities and modes of experience" incorporating "computational sensors that monitor environmental pollution, to organisms that sense and bio-accumulate environmental toxins, and satellite that remotely sense aquifers" (2019, 724).

In this chapter, I want to focus on a particular constellation of sensing entities that together form a kind of "machinic sensibility" (Hong 2016), within the "driving-machine" (Hind 2019) itself. Machinic sensibility, in Hong's definition, describes "technical objects' own ability to sense the material world, and derive information through this process, in ways that are always entangled with, but ultimately distinct from human sensibility" (Hong 2016, 15). Here, media are only "*indirectly* correlated to human modes of experience," in which "the avenue of their impact on human experience and of their implications of humans within their operationality has shifted from a direct to an indirect modality" (Hansen 2015, 6, emphasis in original). In Hong's words, "such engineering *entirely bypasses, occurs prior to, and in sensory regions inaccessible by,* the human subject" (Hong 2016, 15, emphasis in original).

Machinic sensibility, then, is defined by an operational agency in which kinds, or modes, of sensing occur without direct correlation to, or impact on, human experience. Thus, whilst Gabrys (2019) extends the notion of sensing *practices* beyond the strictly human, to all manner of other possible technological and biological agents, both Hong (2016) and Hansen (2015) point towards a different kind of sensing *operation* largely occurring beyond or outside the human, in which to some degree, sensemaking is automated and/or autonomous (Andrejevic and Burdon 2014). Here the point is not that human awareness of, or access to, these sensemaking procedures is entirely impenetrable; but that these sensing processes are functionally distinct and independent from (human) awareness or access. In other words, they do not require direct human involvement to engage in sensemaking activities. This is what Hansen alludes to when he discusses the "veritable inauguration of new, *properly technical* domains of sensation" brought into being through the development of "machinic sensors *that possess sensory domains of their own*" (2015, 54, emphasis added).

I argue that this machinic sensibility is dependent upon four orders, or interpretations of sensing, expanding on Hong's own two-fold distinction. Firstly, this sensibility is a process of *feeling*, in which the likely forms of phenomena are sensed. For lidar, this feeling is enacted at the point of contact between individual pulses of light and objects within the urban environment. Only after the return of many more pulses do such objects start to come into view, with shapes, textures and contours rendered increasingly visible as a lidar unit scans the landscape. Secondly, this sensibility also invariably entails *meaning* making, in which phenomena are made sense of, or understood. Within autonomous vehicles, as I will discuss, this meaning making is distributed, even if lidar is responsible for the bulk of the sensing.

Beyond these two definitions that Hong identifies, I argue that the term machinic sensibility also denotes a process of capture (Agre 1994; Gekker and Hind 2019), in which the form (feeling) and comprehension (meaning) of phenomena are recorded, stored, and utilized in order to enhance the vehicle's ongoing perceptive capabilities. Lastly, this sensibility is meant to arrive at a good decision; that is, a normative outcome deemed "sensible", as it is encoded into decision-making software. This final interpretation posits that sensemaking is not a neutral pursuit, based only on the application of established scientific principles (for example, lidar and the speed of light), or computational limits (image processing times), but guided by expectations, and conventions, on the "social road" (Brown and Laurier 2017).

Automated, or autonomous, sensing operations can thus be said to "broker human accessibility" to the urban environment, with machinic sensibility constituting a different "domain" of sensibility, in which meaning is derived differently (Hansen 2015, 6). This access, I will contend later, is brokered through novel modes of machinic supervision within the autonomous vehicle, as human drivers become expected to monitor, and document, otherwise "autonomous" sensing operations. Expanding on how machinic meaning making is distinctive, Bunz suggests:

> Artificial Intelligence [AI] systems specialized for object recognition in images [...] identify objects depicted in an image in a very particular way: they record the pixel formations i.e., edges and textures of an image, and its shades and different regions of colour, to then calculate statistically the highest possibility [for] what those formations of edges might illustrate. (2019, 272)

In this characterization of AI image recognition processes, AI systems do not interpret images in the same way as humans. Rather than scanning an image for things that we think resemble familiar objects (a human face, a tree, a building), AI systems trained in object recognition instead consider the properties of these objects as they are composed in the image itself. In such systems, Bunz continues, "meaning is not understood but *calculated*" (2019, 272, emphasis added), with meaning derived instead from statistical confidence or likelihood that an object in an image is as it is according to its properties. Thus, that the calculation of such meaning occurs through a kind of *feeling* in which edges, textures and shades become critical sources of information.

It is this calculated form of feeling that guides lidar, with systems capable of measuring the reflectance of surfaces based on the "intensity" of lidar returns. However, lidar's ability to offer such insight is necessarily shaped by the technical limitations of the type or model of lidar device. Typical products used in prototype autonomous vehicles include Velodyne Lidar's Puck and HDL-64E models. The Puck, as the name suggests, is shaped like a hockey puck and has a 100m range, "best-in-class accuracy and calibrated intensity" as well as a "sensor-to-sensor interference mitigation feature" (Velodyne Lidar 2020a). It is commonly used by manufacturers to provide additional lidar sensing support along the side of the vehicle. The HDL-64E, on the other hand, is a "high definition real-time 3D lidar" with an enhanced 120m range, sixty-four channels, a 360° horizontal field-of-view, capable of generating "up to around 2.2 million points per second" (Velodyne Lidar 2020b). It is typically used to provide principal lidar capabilities on the roof of the vehicle (as illustrated in figures 3.1 and 3.3), and can usually be identified by the rotating casing that exposes the sensors whilst in operation.

As a Velodyne Lidar executive has contended, "the resulting point cloud of distance and intensity information is so dense that computer programs can identify objects such as street curbs and overhead wires at distances of over

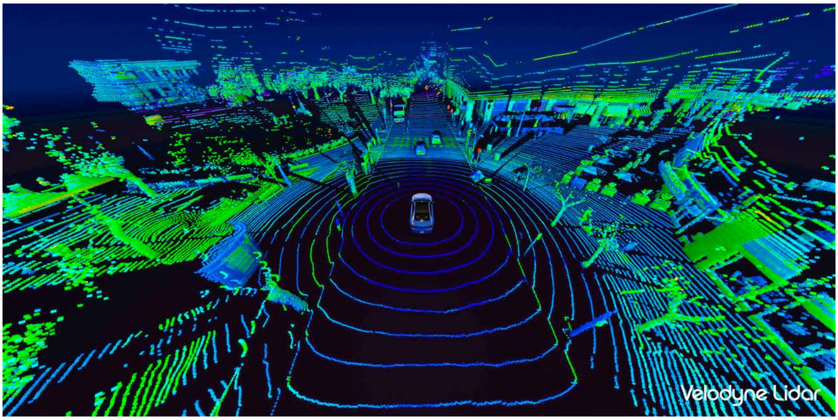

Figure 3.2. A lidar point cloud with return "intensity" visualized in colour. Courtesy of Velodyne Lidar.

100m" (Schwarz 2010, 429). However this claim, of "around 2.2 million points per second" is "configuration dependent" (Velodyne Lidar 2018a, 2). It is this configurative dimension that is central to the sensing capacities of the lidar model in question, allowing it to adapt, or be adapted, to different situations.

The HDL-64E can operate in two modes: single return and dual return. Single return mode only offers a density of around 1.3 million points per second (a less pointy cloud), where the lidar pulse simply records the first thing it hits (i.e. a "single" return). Dual return mode provides the magical figure of 2.2 million points per second, recording multiple hits instead. The latter, therefore, provides an evidently richer account of the urban environment.

On dual return mode, the manufacturer notes that "different environmental conditions require a different priority of the type of distance point returns" (Velodyne Lidar 2018b, 15). For instance, the unit can prioritize the "strongest" distance points (the default). Or, if desired, the last distance point returned can be prioritized. As further suggested, "poor visibility conditions, such as fog and dust, benefit from collecting the distance return values based on the 'last return' scenario". This means that the "near field occluding atmosphere is ignored", i.e. the area containing fog or dust (Velodyne Lidar 2018b, 15). This is another example of where the sensing capacity of the lidar model is configuration dependent. In a last return scenario, these "near things" are deliberately ignored, constructing an image of the urban environment that deliberately discounts the real-world presence of some objects.

Thus, both the *distance* of data collected and the *intensity* of data collected are contingent upon the calibration of the unit itself, radically transforming the ability of the lidar model to feel the urban environment, capture data on the nature of these interactions, derive meaning from them, and ultimately to execute good, or acceptable decisions.

Distributing capacities

However, this machinic sensing is not performed in a singular location, nor executed by a singular entity. Instead, machinic sensibility is dependent on the distribution of sensemaking capacities throughout the vehicle itself. Here I contend that this sensemaking is, firstly, spatially distributed: sensing not only takes place in different locations but is also "oriented" differently towards a surrounding environment. But, secondly, sensemaking is also informationally distributed: sensor data is variously distributed to different parts of the vehicle in order to execute acceptable decisions. In this section I consider how these distributive capacities might be conceived.

As Munster and Lovink (2005) write, "new media are increasingly distributed media", requiring a "distributed aesthetics" that "must deal simultaneously with the dispersed and the situated, with asynchronous production and multi-user access to artifacts [...] on the one hand, and the highly individuated and dispensed allotment of information/media, on the other". Sensemaking in the autonomous vehicle is predicated not only on such a distributed aesthetics, of which the asynchronous production of, and multiuser access to, images is the norm, but also by a distribution of capacities through which images can be produced. Thus, the sensemaking capacities of autonomous vehicles are more than a kind of "distributed cognition" in which "machines [...] operate with an autonomy that underwrites our need to rely on them without understanding them" (Hansen 2009, 310). In other words, the "complex distributions of cognition beyond consciousness" are enabled, but also made complex, by distributed sensemaking (Hansen 2009, 310).

More accurately, sensemaking in the autonomous vehicle is dependent on what Munster and Lovink refer to as "loops of dispersal", in which there is "no singular or 'end use' of/for information but rather the endless relaying of media, practices and experience as successive dispersals" (2005). Whether intentional or not, Munster and Lovink valorise both successive and *successful* loops of dispersal, in which the so-called "endless relaying" of media results in an indeterminable volume of differentiated images. I argue here, however, that whilst distributed sensemaking might embody Munster and Lovink's endless, successful relays, these capacities are perhaps better understood in reference to musical composer William Basinski's *The Disintegration Loops*. A set of ambient productions completed as the 9/11 attacks were happening, the records were made when Basinski attempted to digitize a set of analogue tape loops. Rather than a flawless transfer of original compositions made by Basinski in the 1980s, a series of altogether more ghostly recordings were produced as the metal coating on the tape loops proceeded to blister and physically disintegrate (Richardson 2012).

Sensemaking in the autonomous vehicle is very much dependent on an endless relay of information between sensing units, systems, and other physical components such as brake modules and steering wheels. In other words, loops of dispersal. However, in many situations, these relays do not always work as intended. Instead, they are better characterized as *loops of disintegration* as sensor units are wrongly calibrated, sensor data is poorly captured, objects incorrectly identified, and decisions wrongly executed. Yet rather than bringing these relays to a halt, like Basinski's tapes they generate entirely new forms: new point clouds, new "clusters" of data points (Amoore 2018), new trajectories, and ultimately new decisions.

Yet whilst machinic sensibility is dependent upon a sometimes-disintegrative distribution of capacities throughout the autonomous vehicle, it also engenders a "functional" (Pasquale 2017) or "infrastructural" (Bratton 2015) auto-*nomic* sovereignty (Gekker and Hind 2019) enabled by the reliability, accuracy, and comprehensive qualities of lidar. As Velodyne Lidar contends, using lidar alongside cameras and radar, "allows better field of view and makes more accurate localization and free space detection possible" (Velodyne Lidar 2018c, 6). Moreover, in low light conditions, "lidar significantly fill[s] in the gaps created by the limitations of [...] other sensors" (2018c, 6). In this, lidar's sovereign status is derived from its ability to produce more useful, nominally accurate, data in a variety of situations. The framing of lidar as a sovereign actor is not to suggest it either acts alone, or even acts at every decidable moment. Instead, it is to suggest that as a sovereign actor, other sensing systems work with, for, and under it. Whilst figure 3.3 elides the distributed nature of sensemaking in a prototype autonomous vehicle, it nonetheless illustrates lidar's sovereign status, to which other modes of sensing are typically subordinated. Rather than being non-existent or invisible, as in figure 3.3, these other modes offer critical support for sovereign sensemaking.

The issue of sovereignty and autonomous vehicles has typically been couched in moral terms, most evidently through the "moral machine" project (Awad et al. 2018) and the "trolley problem" (Ganesh 2017), in which decisions around who to "save" and who to "kill" are rendered in utilitarian terms.

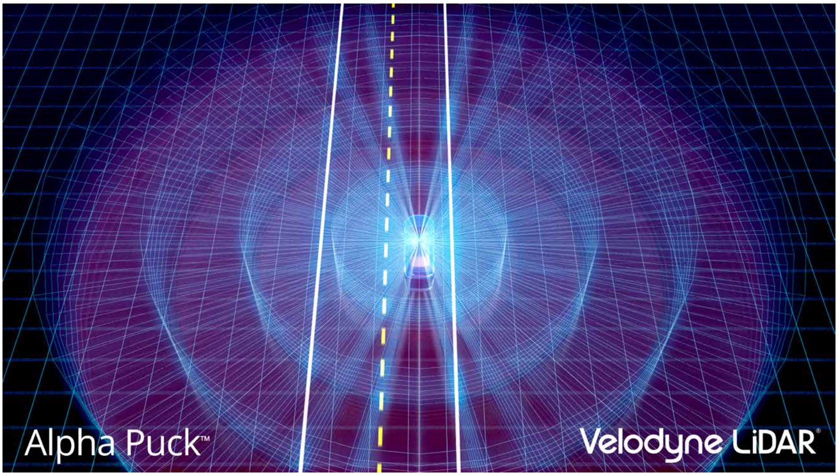

Figure 3.3. A stylised illustration of lidar's "sovereign" sensing capabilities, eliding its distributed nature. Courtesy of Velodyne Lidar.

Yet, limiting the discussion around machinic decision-making to moralistic debates ignores how the technical arrangement of sensing systems and attendant algorithmic software derive or calculate meaning, as discussed earlier. In this, there is no machinic desire to "make moral decisions" (Awad et al. 2018, 1); machines only desire arriving at acceptable decisions as they are calculated by onboard systems.

To consider how this distribution of capacities operates, I will turn for the first time to the Uber crash in Tempe, Arizona in March 2018. Here, I contend that the sovereign status of lidar is best explained in how sensor data captured of the urban environment is used to categorize other road users, as the bounding boxes in figure 3.1 show.

As Elaine Herzberg was walking across Northbound Mill Avenue in Tempe, Arizona, she was detected by an Uber ATG developmental automated driving system (ADS) onboard a modified Volvo XC90 test vehicle. To perceive the surrounding environment, the vehicle was equipped with 20 ultrasonic sensors, ten cameras, eight radar sensors, and one lidar unit (National Transportation Safety Board [NTSB] 2019a, 4). In the 5.6 seconds before Herzberg was hit, she was classified by the ADS on ten separate occasions, with each classification yielding a different possible trajectory Herzberg might take across the road (NTSB 2019a, 10-11).

On the first occasion, Herzberg was detected by the radar system as a *Vehicle*. 0.4 seconds later, she was detected by the lidar system and deemed to be a static object, putting her into the category of *Other*. One second later she is classified again as a *Vehicle*, but nonetheless is still presumed to be static. 2.6 seconds before impact, the ADS reclassifies her for a fourth time; this time as a *Bicycle*, deciding the bicycle by her side is being ridden. With 2.5 seconds left, the system finally predicts she is moving, yet through a lane adjacent to the test vehicle. 1.5 seconds before impact she is again classified as *Other*, and all previous trajectories are "reset". She is once again deemed to be a static object. At 1.2 seconds before impact, she is reclassified for a final time, now as a *Bicycle*, with the ADS predicting she is in the direct path of the test vehicle. Now too late to safely execute an emergency avoidance strategy, the ADS initiates "action suppression" designed merely to mitigate the effects of an impact. 0.2 seconds before Herzberg is hit, action suppression ends and the system issues an auditory warning. 0.02 seconds before impact, the vehicle operator (VO), Rafaela Vasquez, takes control of the steering wheel; now powerless to prevent the fatal crash (2019a, 10-11).

Here, sensemaking capacities are distributed variously. Firstly, through the processes of object detection and classification built into the ADS. With each subsequent classification – first as a *Vehicle*, then as *Other*, finally as a *Bicycle* – these capacities mutate, rendering Herzberg in different terms on each occasion. Secondly, between sensing systems in the vehicle itself, most notably between the radar system that first identifies Herzberg, and the lidar system that subsequently classifies, then reclassifies, her. In this, whilst the radar system is the first to pick Herzberg up, with its superior range detection, it is lidar that ultimately takes over as the vehicle approaches her. Thirdly, and belatedly, sensemaking capacities are distributed between the vehicle's sensing systems and the physical components designed to prevent a collision, such as the brakes or steering wheel. With this, the ADS communicates its decision, principally reliant upon the erroneous classifications based on lidar data, to the relevant components designed to perform the necessary actions. Then lastly, and even more belatedly, stepping outside of the intended, idealized, closed integration loop between these various sensing systems and physical components: the human VO herself contributes to the sensemaking capacities of the autonomous vehicle. Across these many capacities, sensemaking is not only distributed imperfectly, but catastrophically.

Here it becomes obvious that the vehicle in question did not, and was not, simply making a single moral decision at a nominal crossroads like in the fabled trolley problem. Instead, the system was engaged in an ongoing assessment of criteria, evaluating Herzberg at various stages, categorizing her differently each time, and making ongoing decisions to act (or not) on each occasion. At each stage, a different snapshot of the urban environment is made, with sensor data used to calculate the meaning of the objects in view. In sorting Herzberg into different categories the vehicle was reliant on the sovereign qualities of lidar. The tragic conclusion that can be drawn from this was that Herzberg was not moving "properly" or "normally" enough, or indeed, not moving "in the right place" within the urban environment, to be made sense of.

Distancing sense, prioritizing "recency"

Autonomous vehicles are being "computationally optimized for terrains [...] incorporating the sensing of elemental, atmospheric, and meteorological phenomena" (Hind 2019, 402). Consequently, as Gabrys and Pritchard (2018) argue, "distinct affective and political capacities are operationalized through [such] sensing practices". I want to argue here that not only are *distinct* capacities operationalized through the sensing operations of the autonomous vehicle, as articulated in the previous section. But in addition, that *distance* – most notably, the distance between vehicle and object(s) – becomes a significant variable in how these capacities are operationalized, as made evident in the death of Elaine Herzberg.

Ash argues, in reference to the Tesla Model S, that it is unhelpful to "understand smart objects' sensory capacities in the form of metrical distance" (2018, 170), despite it being used to promote the vehicle's "autopilot" driver-assist feature. Ash contends that such systems should be "defined by their capacity to *differentiate* between objects and *assign* the correct references to [...] objects to make distance sensible and intelligible" (2018, 170, emphases added). Metrical distance alone is no measure of the "smartness" of an object, nor indeed, of its sensemaking capacities. As Ash reiterates, "it does not matter how 'far' a sensor can reach, if that sensor cannot differentiate between objects [...] and so enable a car or driver to assign the correct references to those objects" (2018, 170).

To add to Ash's analysis, it is important to recognize that whilst the "smartness" of an object is not built (only) on its depth perception, neither is it based on universal perception. Autonomous vehicles are often touted as having "360 degree view" (Oxbotica 2019), or that specific systems can provide "360° [...] coverage" (NTSB 2019a, 4), or can "detect objects in a 360-degree area" (NTSB 2019a, 5), as illustrated in figure 3.3. In these statements, distance is mobilized differently, as a capacity of the vehicle to offer comprehensive depth perception. What these claims elide, however, is not only the composite nature of this apparently seamless and "universal" perception, but also the varying perceptive depth offered in 360 degrees. In other words, purported 360 degree vision is offered only through the integration of multiple units with specific sensing capacities, which in doing so, create an uneven depth to this purported capacity. Some sensing units may offer greater depth (radar) than others (ultrasonic sensors), whilst some may necessarily overlap (forward cameras and lidar) whilst rendering distance differently (compare radar and lidar).

Thus, it is only through a technical comprehension of distance that object-recognition, and therefore object differentiation, occurs. In the case of Uber, this is made possible through what it calls a "prioritization schema" that promotes "tracking by certain sensory systems over others" (NTSB 2019d, 12). Such a schema is "also dependent on the recency of an observation", where recency is defined as the "more recent detection of an object" (NTSB 2019d, 12). In other words, that some sensing systems, and some detection events, are prioritized over others at any one time. This is whilst lidar units, such as the Velodyne Lidar Puck or HDL-64E models discussed before, are also calibrated to prioritize either the strongest or last distance point recorded. An acknowledgement of the contestability of such a schema was made by Uber, post-crash, when it announced it would change the way the system "fuses sensor information" when predicting object trajectories (NTSB 2019d, 13). In any case, both distance point prioritization and sensor system prioritization are critical features of the prototype autonomous vehicle.

Take, once again, the moments before Herzberg was hit. 5.6 seconds before impact, she is first detected by the vehicle's radar system. Two radar units provide forward scanning and can operate in two modes. Mode one, a long-range scan, has "an observational range of up to 180 meters with a 20-degree field of view", whilst mode two, a medium-range scan, has "an observational range of up to 65 meters with a 90-degree field of view" (NTSB 2019a, 5). As the report continues, the "radar processing units conduct the initial processing of the [sensed] data, which the ADS then uses to build and continually update the representation of the surrounding environment" (2019a, 5). Whilst it is unclear which mode was active at the time, Herzberg was recognized as a vehicle. Thus, at 5.6 seconds before impact, Herzberg's distant capacities are deemed to resemble a vehicle; likely because she is simply present in a vehicle lane. Nevertheless, mere (metrical) distance is enough for such a recognition to occur; distant capacities are operationalized through the sensing operations of the vehicle. Metrical distance matters because, computationally and operationally, the radar unit attached to the vehicle has a sensory limit; either up to 180 metres, or 65 metres, depending on the operative mode.

Yet in this integrated process, as contended, some sensor systems take priority. At the time of the crash, only the Uber ADS was active. However, the Volvo XC90 was also equipped with a parallel advanced driver assistance system (ADAS) called City Safety. Although not a fully automated driving system, City Safety is designed to detect pedestrians in urban environments; comprised of what Volvo calls Forward Collision Warning and Automatic Emergency Braking. When the vehicle was being used in manual mode, controlled by a VO, "all the Volvo ADAS components were active and operated as designed" (NTSB 2019a, 13). Yet when the Uber ADS was activated, "all Volvo ADS components were automatically disengaged" (NTSB 2019a, 13). Only the vehicle's passive safety technologies, such as seatbelt pretensioners and airbag deployment systems, "remained active" in autonomous mode (NTSB 2019a, 14).

Two reasons are given for why the Volvo system was deactivated at the time of the crash. Firstly, that because the Uber ADS and Volvo ADAS both used radar, there was a "high likelihood of misinterpretation of signals" (NTSB 2019a, 14) between both. Secondly, that in receiving braking commands from either system, the "vehicle's brake module [would] not [have] been designed to assign priority" to either system (NTSB 2019a, 14). Subsequently, two sets of unresolvable conflicts occur.

Firstly, there is an identified or presumed conflict between sensing approaches. Here the issue is not that each individual system uses different sensing methods (one using lidar, the other radar, for instance), but that both use the same approach, i.e. radar. Likely due to respective system configurations, radar data will be processed and made sense of differently by each system. The result is differently interpreted data of the same phenomena using the same method. Secondly, there is a conflict between composite automation/assist systems. Here the issue is that each individual system – Uber's ADS and Volvo's ADAS – will likely send similar commands to the various modules in the vehicle assigned to move physical components such as the brakes. The result is possibly conflicting commands issued to components not programmed to decide which to listen to or ignore.

Ultimately, this means some sensing units, and some composite systems, as well as some detection events, are prioritized over others. The consequence of these conflicts – presumed or actively identified – is that some modes of distancing are prioritized over others; meaning only some distant capacities are operationalized at any one time. Why this matters is that the capacities of other road users in the urban environment are only realized through some sensing systems, and those identified more recently assume greater priority. Understanding when and where particular modes are themselves prioritized is critical to articulating the effect of these sensing systems on how the urban environment is variously perceived at any one time, according to the registered, and classified, capacities of other road users.

Enabling machinic sensibility, or "what's in a glance?"

The result of both a distribution of capacities and a prioritization of sensing is a differentiation in experiential effects. In arguing that machinic modes of sensing constitute a different "domain" of sensing (Hansen 2015, 6), I have not intended to erase the involvement of human actors in the operation-at-large. Instead, as outlined before, I argue that these sensing operations "broker human accessibility" (Hansen 2015, 6) to the urban environment. As the first section of this chapter hinted at, human actors in such arrangements become supervisors, overseeing how the machine operates. This was a role performed by Rafaela Vasquez in the fatal Uber crash in Tempe, Arizona, but also by many other VOs employed by the company as nominal machinic supervisors. In this final section I want to draw attention to the specific experiences of Rafaela Vasquez as affected by the distribution of capacities at the time of the crash: both subject to, and an unwitting enabler of, machinic sensibility. In other words, the sensing operations of the Volvo XC90, equipped as it was with an in-development Uber ADS, were only made possible through the interventions, interpretations, and interactions of human operators like Vasquez – or, indeed, the lack thereof.

Firstly, as a VO, Vasquez was responsible for carrying out a range of tasks before, during, and after testing. When the vehicle was in autonomous mode, she would have been expected to do three things: (a) continuously monitor the state of the vehicle and the road (b) take control of the vehicle should a dangerous situation arise, and (c) document performance-related incidents. In order to train VOs to perform these tasks correctly, they are subject to a three-week "onboarding process" in multiple locations, where they are taught vehicle handling skills, and introduced to various scenarios to "test [...] [their] decision making skills and ability to interact with the vehicle controls" (NTSB 2019b, 3). Then, VOs are tested on company procedures and processes, before being "re-localized" in relation to state driving laws in Arizona, and introduced to Uber ATG's infraction policies and test routes. Although Vasquez completed the training in a slightly different order, she followed the same three-week training course, intended to equip her with the skills to be a VO.

Yet Vasquez was originally trained on passenger operations (as opposed to test operation) according to a pilot/co-pilot model. In this format, two VOs would be present in any one test vehicle. One VO would occupy the driver's seat, ready to take control if a situation arose. The other VO would occupy the front passenger seat, supervising the vehicle's path, whilst tagging and annotating issues on a laptop that might arise whilst the vehicle was in autonomous mode. In this configuration, the three principle tasks for each VO, as outlined above, would have been divided between two VOs: VO1 (pilot) principally responsible for (a) and (b), whilst VO2 (co-pilot) principally responsible for (c). However, in October 2017, things changed. As the report details:

> Uber ATG integrated much of the co-pilot's functions into the 'front seat control application' (FSCA) software, housed on a centre-dash mounted tablet computer in the SDV [self-driving vehicle]. The FSCA interface was the primary means for the VO to interface with the SDS [self-driving system]. Complex functions on the FSCA were locked out once the SDV was in motion, and according to Uber ATG, functions that were available to the VO while the vehicle was in motion only required one to two taps to complete. (NTSB 2019b, 3)

In short, Uber consolidated the role of pilot and co-pilot into one VO and the aforementioned FSCA software. The result was that tasks (a), (b) and (c) – continuous monitoring, possible control, and performance documentation – were now expected to be performed by a single VO, sitting in the driver's seat. Not long after, Vasquez was trained on the interface, beginning work as a single VO a month later. The previously distinct training paths of passenger operations and test operations were now combined to reflect these changes.

Thus, Vasquez and all other VOs were responsible for interacting with FSCA software on tablet, affixed to the centre dashboard of the vehicle. Moreover, VOs were still expected to complete interactive tasks while the vehicle was in motion. Whilst, as the excerpt above mentions, "complex functions" were "locked out" whilst on the move, VOs were still required to perform other functions requiring "one to two taps to complete" (NTSB 2019b, 3). The report details four such input types, including "tagging an object of interest", "notifying the engineering team of an on-vehicle issue", "tagging incidents or infractions" and "tagging when the SDS performs incorrectly" (NTSB 2019b, 8). Thus, whilst each function might only have required one or two taps, the combined occurrence of these problems could demand repeated interactions with the tablet.

These functions were visually represented on the interface itself. If a VO wanted to tag an object of interest, they could locate the "label" icon in the bottom-left corner of the screen. If there was an on-vehicle issue, the VO could tap the "ticket" icon at the bottom-centre of the screen. If the vehicle had been involved in an incident or infraction, the VO could tap the "attn" (attention) icon, again, alongside the ticket option. If the autonomous system had acted strangely (although not necessarily dangerously), then the VO could press the "autonomy" icon at the bottom-right corner of the screen. Thus, in order for the VO to perform their ordinary duties – namely, the documentation of vehicle performance – they would have to get used to tapping the dashboard-mounted interface whenever necessary. All logged incidents would then be dealt with by relevant ATG teams, responsible for fixing or updating the responsible features. Test iterations – and, specifically, the documentation of incidents during them – were critical stages in the development of the sensemaking capacities of the Uber autonomous vehicle. Without the recording of these incidents – possibly unencountered in other test modes or simulated situations – the vehicle system might well be worse at making decisions, recognizing other road users, or obeying local traffic laws.

For the VOs like Vasquez, attention would naturally be divided between road and interface, windscreen and dashboard. In the final report published after an eighteen-month investigation, the probable cause was given as "the failure of the vehicle operator to monitor the driving environment and the operation of the automated driving system because she was visually distracted throughout the trip by her personal cell phone" (NTSB 2019d, 59). In records obtained from video streaming providers (including Hulu), NTSB determined Vasquez "was continually streaming a television show between 9.16pm and 9.59pm […] That period covered the entire crash trip, which included 39 minutes on a public road" (NTSB 2019d, 24). These conclusions were drawn despite Vasquez stating she had "placed her personal phone in her purse before driving, and that her company phone was on the passenger seat at the time of the crash" (NTSB 2019d, 24).

Here, the intention is to not disagree with the conclusions drawn by the NTSB about the crash, after which Vasquez was charged with negligent homicide (Levin 2020). Nor is it to believe Vasquez's account of the crash; that her personal phone was in her bag, placed on the back seat of the vehicle, both out of sight and out of reach. Rather, the intention is to make sense of the tasks required to be performed by any VO whilst the vehicle is in autonomous mode, and those not permitted, i.e. like using a personal mobile phone. In other words, this chapter seeks to identify the precise role of – and the specific risks taken by – a VO ultimately responsible for enabling the eventual sensemaking capacities of the autonomous vehicle.

As an interview with Vasquez suggests, the latest VO training "indicated that she [VOs] may look at the iPad for 5 seconds and spend 3 seconds tagging and labelling" (NTSB 2019c, 6). VOs were expected to look forward at all times, including (indeed, especially) when the vehicle was in autonomous mode. Yet, they were also expected to perform tagging and labelling tasks as regularly as required, with up to 8 seconds spent looking at, and interacting with, the central dash-mounted tablet. As interior photos show, the lower console area "where a cell phone could be placed" (NTSB 2019b, 7) was directly underneath where the tablet was mounted. The NTSB deduced:

> From the time the VO exited the parking lot to the time of the crash, the VO frequently glanced down towards the lower centre console area. The Tempe Police tabulated the number of glances the VO made towards the lower centre console area during a 27-minute window, from 9.31pm to 9.58pm. During this timeframe, the VO glanced down at the same spot 204 times, of which 166 instances were when the vehicle was in motion. The[y also] estimated that […] the VO's eyes were averted from the roadway [for] approximately 32% of the time. (2019b, 7).

Much meaning is attributed to the "glances" made by Vasquez towards the lower console area, and the frequency at which these glances occurred during the time the vehicle was in autonomous mode. Yet glancing towards this area was not against Uber policy. Indeed, as has been suggested, it was part of the assumed role of any VO – to look towards, and interact with, a tablet mounted on the central dashboard whenever an incident arose that required documenting. Necessarily, in doing so, VOs would have to look away from the road ahead, and down towards the interior of the vehicle; as well as concentrating on making an accurate record of any encountered incident.

Thus, this shift in attention was part of Vasquez's – and any VOs – assumed responsibilities. Without taking such action – repeated glances, diverted attention, concentration, tapping, and tagging – the developmental Uber vehicle would be without critical operational insights derived from test situations. In other words, the vehicle would likely fall short – just like it did in this crash – of correctly sensing other road users, and adapting to their presence. The future sensemaking capacities of the autonomous vehicle being tested were dependent on routine glances, just not the kind Vasquez was deduced to have made.

Conclusion

In this chapter, I have argued that whilst lidar is central to the sensemaking capacities of prototype autonomous vehicles, this sensemaking is only made possible through the distribution of responsibilities throughout any such vehicle. Further, I have contended here that this sensemaking is only enabled through the involvement of human operators involved also in the correction, and verification, of machine-readable driving worlds. This capacity is what Hong (2016) refers to as "machinic sensibility", a process through which technical objects recognize things in the world, and derive information from this recognition. Importantly, machinic sensibility is entangled with other forms of human sensing, visual and otherwise – whether in the form of quality control, oversight, or decision-making. Nevertheless, this machinic sensibility is better characterized through the figure of the sensing operation held at arms-length from human intervention.

In this, I have suggested that the machinic sensibility of lidar in the prototype autonomous vehicle is dependent upon four orders of sensing. Firstly, through a process of *feeling* or the interpretation of the shape and form of phenomena. Secondly, and necessarily, through a process of recording and *capturing* such phenomena, so that this feeling can be made operational. Thirdly, enabling the processual making of *meaning* through which phenomena are "made sense of". Then, lastly, through the execution of *good* decisions – a normatively-derived outcome deemed "sensible" and reasonable to at least some of the involved parties.

Yet, the machinic sensibility of lidar in the prototype autonomous vehicle is not being singularly, and solely, performed by and in the lidar unit itself. Instead, this machinic sensibility is dependent on the distribution of sensemaking capacities throughout the vehicle. This, I have argued, involves both a *spatial* distribution between components capable of aiding the four orders of sensing (feeling, capturing, meaning, good) and an *informational* distribution in which data is variously distributed to enable the smooth execution of decisions. Sensing is distributed to verify and authenticate sovereignty, exemplifying a case of functional or infrastructural auto-nomic sovereignty.

This machinic sensibility, however, is also dependent upon the operationalization of distant capacities. In this, the nominal distance between any lidar-equipped vehicle and objects within the urban environment is a critical factor in their being sensed. This operationalization is referred to as a "prioritization schema" (NTSB 2019d, 12) in which objects closer to the vehicle are prioritized over those further away. Moreover, "recency" (NTSB

2019, 12) – or the more recent detection of an object – is given priority over objects sensed longer ago.

The distribution, and distance, of machinic sensibility is, I argue, dependent on its enabling. Here, under specific test conditions, machinic sensibility as an operation is surfaced, or made available to human operators. In such instances, these human operators – and the tasks they are required to perform – are not only actively shaped by the operational capacities of lidar, but also the various interfaces that allow them to interrogate these capacities during test situations. As such, I contend that this surfacing, or availability, structures and scripts the experience of those made responsible for fine-tuning the sovereign sensemaking capacities of autonomous vehicles.

Throughout this chapter I have drawn on both off-the-shelf lidar products, as well as the specific testing of developmental autonomous vehicle systems. Most notably, I have focused on the crash in March 2018 in Tempe, Arizona, involving a prototype autonomous vehicle, that killed Elaine Herzberg. In the first instance I have suggested that Herzberg was subject to the ongoing assessment of operational criteria that led to her being classified, and re-classified, as various objects – from a car to a bike – in the seconds before impact. In the second instance, I have argued that this ongoing assessment was dependent upon her own "distant capacities", being variously sensed by lidar and other perceptive systems in the prototype autonomous vehicle, at different times. In this, Herzberg was interpreted, captured, made sense of, and ultimately decided on differently, at different distances to the vehicle itself. Then, thirdly, I moved on to Rafaela Vasquez, the nominal operator of the prototype vehicle involved in the crash itself. Here, I contended that her role as a diagnostician of the sensemaking capacities of the vehicle led to scrutiny of the application of her tasks as a certified vehicle operator. In this, I have queried the significance of the "glance": the repeated actions Vasquez is alleged to have made that impaired her ability to take control of the vehicle in the seconds before the crash. The sensemaking capacities of these prototype autonomous vehicles are dependent upon the interpretive, and interactive, work of vehicle operators such as Rafaela Vasquez.

What I have sought to do in this chapter is to give colour to the sensing operations performed by a prototype autonomous vehicle, particularly to how it perceives urban space, and to highlight bundled processes and practices that coalesce around these operations. What is critical to note, therefore, is that the chapter has not speculated on any eventual or hypothetical sensemaking capacities of an autonomous vehicle. Instead, it has sought to articulate the sensing operations of prototype autonomous vehicles being tested at this moment, to make sense of how these operations are not only being performed, but also necessarily upgraded and improved. As such, it is a snapshot of the sensemaking capacities of "more-than-lidar" and the various loops of interpretation, meaning-making, and decision-making that comprise this arrangement. The perception of urban space – including perceiving it visually – is enabled or indeed disabled through these loops, in which particular objects are sensed, and made sense of, at any one time. When these loops short-circuit or disintegrate, as was the case in the Uber crash, sensemaking does not stop. Instead, novel, unintended, and potentially catastrophic effects result, generating a differentiation of experience, whether for other road users or those responsible for supervising the work of machines.

References

Agre, Phil. 1994. "Surveillance and Capture: Two Models of Privacy." *The Information Society* 10 (2): 101–27.

Amoore, Louise. 2018. "Cloud Geographies: Computing, Data, Sovereignty." *Progress in Human Geography* 42 (1): 4–24.

Andrejevic, Mark, and Mark Burdon. 2014. "Defining the Sensor Society." *Television & New Media* 16 (1): 19–36.

Ash, James. 2018. *Phase Media: Space, Time and the Politics of Smart Objects*. London: Bloomsbury.

Awad, Edmond, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. 2018. "The Moral Machine Experiment." *Nature* 563: 59–64.

Bratton, Benjamin. 2015. *The Stack: On Software and Sovereignty*. Cambridge, MA: MIT Press.

Brown, Barry, and Eric Laurier. 2017. "The Trouble With Autopilots: Assisted and Autonomous Driving on the Social Road." *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, May: 416–29.

Bunz, Mercedes. 2019. "The Calculation Of Meaning: On the Misunderstanding Of New Artificial Intelligence As Culture." *Culture, Theory and Critique* 60 (3–4): 264–78.

Davies, Alex. 2018. "How Do Self-Driving Cars See? (And How Do They See Me?)." *Wired*, August 11, 2018.

Gabrys, Jennifer. 2019. "Sensors And Sensing Practices: Reworking Experience Across Entities, Environments, And Technologies." *Science, Technology, & Human Values* 44 (5): 723–36.

Gabrys, Jennifer, and Helen Pritchard. 2018. "Sensing Practices." In *Posthuman Glossary*, edited by Rosi Braidotti, and Maria Hlavajova, 394–96. London: Bloomsbury.

Ganesh, Maya Indira. 2017. "Entanglement: Machine Learning and Human Ethics In Driver-less Car Crashes." *A Peer–Reviewed Journal About* (*APRJA*) 6 (1): 76–87.

Gekker, Alex and Sam Hind. 2019. "Infrastructural Surveillance." *New Media & Society* 22 (8): 1414–36.

Hansen, Mark. 2009. "Living (with) Technical Time: From Media Surrogacy to Distributed Cognition." *Theory, Culture & Society* 26 (2–3): 294–315.

Hansen, Mark. 2015. *Feed-Forward: On the Future of Twenty-First-Century Media*. Chicago: The University of Chicago Press.

Hind, Sam. 2019. "Digital Navigation and the Driving-Machine: Supervision, Calculation, Optimization, And Recognition." *Mobilities* 14 (4): 401–17.

Hong, Sun–ha. 2016. "Data's Intimacy: Machinic sensibility and the Quantified Self." *Communication +1* 5 (3): 1–36.

Levin, Sam. 2020. "Safety Driver Charged In 2018 Incident Where Self-Driving Uber Car Killed a Woman." *The Guardian*, September 16, 2020.

Lovink, Geert, and Ned Rossiter. 2019. "Propositions on the Organizational Form." In *Organize*, edited by Timon Beyes, Lisa Conrad, and Reinhold Martin, 89–101. Lüneberg: Meson Press.

McCosker, Anthony, and Rowan Wilken. 2020. *Automating Vision: The Social Impact of the New Camera Consciousness*. New York: Routledge.

Metz, Cade. 2018. "How Driverless Cars See the World Around Them." *The New York Times*, March 19, 2018. https://www.nytimes.com/2018/03/19/technology/how-driverless-cars-work.html/.

Munster, Anna, and Geert Lovink. 2005. "Theses On Distributed Aesthetics. Or, What A Network Is Not." *The Fibreculture Journal* 7. http://seven.fibreculturejournal.org/fcj-040-theses-on-distributed-aesthetics-or-what-a-network-is-not/.

National Transportation Safety Board (NTSB). 2019a. "Highway investigation HWY18MH010: Vehicle Automation Report." *NTSB*, May 11, 2019. https://data.ntsb.gov/Docket/Document/docBLOB?ID=40477717&FileExtension=.PDF&FileName=Vehicle%20Automation%20Report-Master.PDF.

National Transportation Safety Board (NTSB). 2019b. "Highway investigation HWY18MH010: Human Performance Factual Report." *NTSB*, May 11, 2019. https://data.ntsb.gov/Docket/Document/docBLOB?ID=40477743&FileExtension=.PDF&FileName=Human%20Performance%20Factual-Master.PDF.

National Transportation Safety Board (NTSB). 2019c. "Highway investigation HWY18MH010: Uber ATG Vehicle Operator Interview." *NTSB*, May 11, 2019. https://data.ntsb.gov/Docket/Document/docBLOB?ID=40477745&FileExtension=.PDF&FileName=Human%20Performance%20Attachment%20-%20Uber%20ATG%20Vehicle%20Operator%20Interview-Master.PDF.

National Transportation Safety Board (NTSB). 2019d. "Highway Investigation HWY18MH010: Collision Between Vehicle Controlled By Developmental Driving System and Pedestrian, Tempe, Arizona, March 18, 2018." *NTSB*, May 11, 2019. https://data.ntsb.gov/Docket/Document/docBLOB?ID=40479021&FileExtension=.PDF&FileName=NTSB%20-%20Adopted%20Board%20Report%20HAR-19%2F03-Master.PDF.

Oxbotica (@Oxbotica). 2019. "Our Eyes Are A Vital Asset When It Comes To Spotting And Reacting To Danger On The Roads." *Twitter*, November 8, 2019. https://twitter.com/oxbotica/status/1192815056926457858.

Pasquale, Frank. 2017. "From Territorial to Functional Sovereignty." *Law and Political Economy Blog*, June 12, 2017. https://lpeblog.org/2017/12/06/from-territorial-to-functional-sovereignty-the-case-of-amazon/

Richardson, Mark. 2012. "The Disintegration Loops." *Pitchfork*, November 19, 2019. https://pitchfork.com/reviews/albums/17064-the-disintegration-loops/.

Rose, Gillian. 2017. "Posthuman Agency In The Digitally Mediated City: Exteriorization, Individuation, Reinvention." *Annals of the American Association of Geographers* 107 (4): 779–93.

Schwarz, Brent. 2010. "Mapping the World In 3D." *Nature Photon* 4: 429–30.

Stilgoe, Jack. 2017. "Seeing like a Tesla: How Can We Anticipate Self-Driving Worlds?" *Glocalism: Journal of Culture, Politics and Innovation* 3: 1–20.

Velodyne Lidar. 2018a. "HDL-64E Data Sheet." HDL-64E. https://velodynelidar.com/products/hdl-64e/.

Velodyne Lidar. 2018b. "HDL-64E User's Manual and Programming Guide." HDL-64E. https://velodynelidar.com/products/hdl-64e/.

Velodyne Lidar. 2018c. "A Safety-First Approach to Developing and Marketing Driver Assistance Technology." *Velodyne Lidar*, December 1, 2018. https://velodynelidar.com/wp-content/uploads/2019/08/ADAS-Whitepaperver.12.2018.pdf.

Velodyne Lidar. 2020a. "Puck: Compact, Powerful, Intelligence." https://velodynelidar.com/products/puck/.

Velodyne Lidar. 2020b. "HDL-64E: High Definition Real-Time 3D Lidar." https://velodynelidar.com/products/hdl-64e/.

About the author

Sam Hind is Research Associate in SFB1187 Media of Cooperation at the University of Siegen. He is co-editor of *Time for Mapping: Cartographic Temporalities* (Manchester University Press, 2018) and co-author of *Playful Mapping in the Digital Age* (Institute for Network Cultures, 2016). He has published in *Political Geography*, *Mobilities*, and *New Media & Society*.