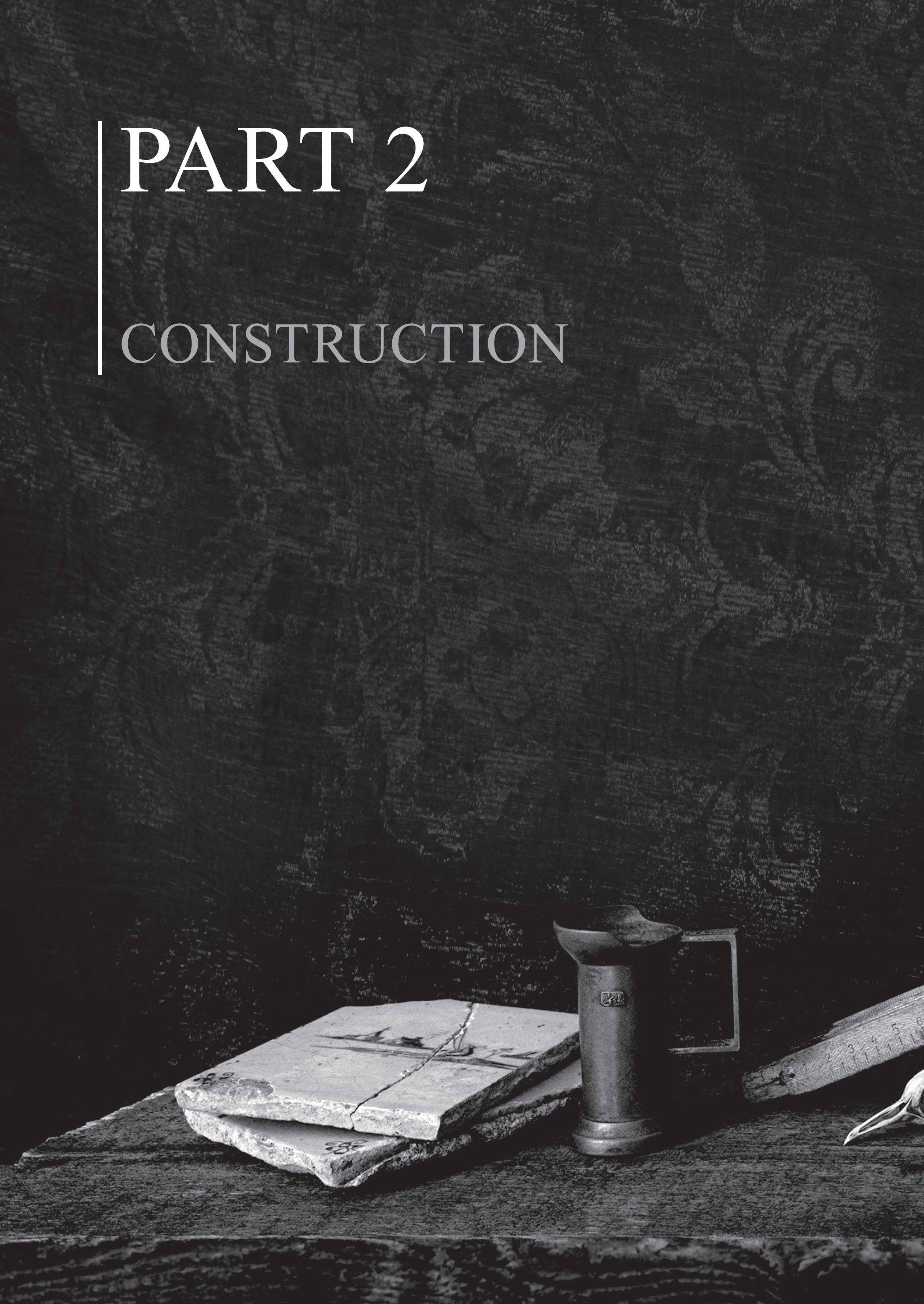
Universiteit Leiden
The Netherlands

Computer says no: understanding digital authority

Weggemans, D.J.

PART 2

CONSTRUCTION



4 RELATION

In 2015, two days before Christmas, there was a press conference in Washington. Governor Jay Inslee told the public that there had been "serious errors with serious implications". He revealed that over the course of thirteen years, thousands of inmates were mistakenly released early from Washington state prisons because of a "computational issue" (Washington governor, 2015). The problem was discovered by the family of an assault victim who had noticed that the perpetrator was about to get out of prison before the date the court had ordered. After they had contacted the Department of Corrections, it was discovered that for more than a decade, the department's software had been erroneously calculating the prison sentences of certain inmates.

It would take another three years before the public learnt about the situation. Meanwhile, more than 3,200 prisoners had been mistakenly released. In some cases, the consequences were serious. One of them, for example, committed an attempted robbery leading to the death of a 17-year-old. Another inmate became suspected of vehicular homicide and felony hit and run. High-ranking state officials had to apologise for the fact that this had happened while the suspects should have still been under lock and key.

The controversy in Washington caused widespread outrage and the agencies involved faced claims from the victim's families as well as from prematurely released offenders themselves who were taken back into custody years later to serve the remaining time (AP, 2018). In front of a state Senate's committee hearing, Dan Pacholke, the Secretary of the Washington Department of Corrections, was asked why the computer's sentence calculations were never checked by human officials. "I don't believe there's a performance system in place that is going to catch a mistake in a computer algorithm", he responded. The chair of the committee commented that he refused to accept this explanation and stated: "this wasn't a computer mistake. This was a people mistake" (Padden, 2016). Pacholke stepped down in the weeks after the hearing. In his letter of resignation, he again apologised for the "tragic consequences of this error" (Pacholke, 2016). At that time, Governor Inslee had already announced a software fix for the sentencing calculations. Until then, however, he insisted that all the upcoming release dates were to be calculated by hand.[64]

This painful story reveals people's willingness to rely on the communications issued by a computer. For years, the computer's output directed the actions of the personnel in Washington State prisons. In the previous chapter it was already stated that the literature on authority almost exclusively focuses on the relations *between people*. This chapter is about the relationship between humans and technologies. Can authority emerge here, and if so, how? Paragraph 4.1 discusses the work of North American philosopher of technology Don Ihde to understand the different ways individuals can relate to technology. In paragraph 4.2, I argue that digital authority represents a newly emerged relationship between humans and technology. I build on this observation and describe the capacity to communicate as a necessary condition for authority. In the previous chapter I concluded that authority always involves acts of communication. To be consequential, technologies thus need to communicate, and this ability is further discussed in paragraph 4.3. The subsequent paragraphs focus on the asymmetry characteristic of authority relationships and the potentialities of persuasion and coercion as constitutive parts of this. Paragraph 4.4 introduces the case of parking enforcement to illustrate the posited asymmetry between digital technology and its users. Paragraph 4.5 discusses how the implementation of digital technologies can severely limit individuals' opportunities for knowledge acquisition. In paragraph 4.6, in turn, I describe how digital authority depends on presumptions of differentiated access and the latent presence of persuasion in the background of the human-technology relation. Paragraph 4.7 considers the potentiality of coercion as a constitutive part of digital authority.

4.1 RELATING TO TECHNOLOGY

Intuitively, digital technologies and other technological artifacts are passive tools, not authorities. They are there to be used and to fulfill specific functions. Philosophers of technology, however, have already compellingly highlighted that technological artifacts cannot be well understood when approached as mere means or instruments to achieve our goals. Instead, the North American Philosopher of Science and Technology Don Ihde (1990) developed a (post)phenomenological understanding of human-technology *relations*, making visible how technologies co-shape human experience and action in the world. In his major work *Technology and the Lifeworld*, he distinguishes four different relationships human beings can have with technological artifacts. First, our involvement with the world can be *mediated* by a technological artifact (Ihde, 1990: 72-97). Ihde describes two forms of this technological mediation. Some technological artifacts merge with our bodies and from this position influence the way we experience our surroundings. Following Maurice Merleau-Ponty, Ihde calls this the *embodiment* relation. Examples of artifacts that form this kind of relation with humans are hearing aids, glasses, canes, and telescopes. When in use, these artifacts withdraw from our attention - they become perceptually transparent and the world is perceived *through* them, accordingly (Ihde, 1990: 72, 86). Ihde schematises embodiment relations as:

Embodiment relations: (Human – technology) → World

In the other structure of mediation, the *hermeneutic* relation, an artifact is also mediating our perception, but the artifact is non-transparent (Verbeek, 2005: 126). Rather than simply withdrawing from our experience, the artifact provides us with a representation of the world that requires our interpretation. The term hermeneutics comes from the Greek verb *hermeneuein,* meaning "to interpret". Ihde uses the thermometer as an example. When you are in your home and you read the thermometer that is nailed to a fence outside, you *hermeneutically* know how hot or cold it is (Ihde, 1990: 84-85). In such situations, we experience reality *with* technology. Ihde schematises hermeneutic relations as:

Hermeneutic relations: Human → (technology - World)

A second set of human-technology relations Ihde identifies is *alterity relations*. Contrary to the previous relations, in alterity relations human beings are consciously involved with a technological artifact. It is the experience of a direct relation *to* a technological artifact (Verbeek, 2005: 126-7). The term alterity is rooted in the work of the French philosopher Emmanuel Levinas who refers to it as "the radical difference posed to any human by another human, an *other*" (c.f. Ihde, 1990: 98). This experience of *otherness*, Ihde says, we can also have with technological artifacts. Alterity relations can emerge for example when people use an ATM, encounter a robot, or prompt a chatbot - and with many other digital technologies. Ihde explains that technologies are experienced as others in alterity relations because they are seen to possess some sort of independence as well as give rise to an interaction between people and technologies (Ihde, 1990: 97 – 108). In alterity relations, as he puts it, there is always a "sense of interacting with something other than me (..), there is a kind of dialogue or exchange" with a technology (1990: 100). We engage in alterity relations with technologies because they appear to be autonomous – these technologies seem to have a life of their own (ibid.). As Peter-Paul Verbeek describes elsewhere:

> "Many toys, such as tops and music boxes, are fascinating precisely because of the apparent autonomy that they possess. Robots possess such an autonomy, to the extent that one could truly speak of "interacting" with these technological "beings". And today, automatic train ticket machines can not only take money and dispense tickets, but also give advice, provide route information, answer questions, and protest when something is done incorrectly" (Verbeek, 2005: 131).

Ihde schematises alterity relations as:

Alterity relations: Human → technology- (-World)

At the same time, Ihde observes that technological artifacts are, at best, *quasi-others*. He describes such otherness as "stronger than mere objectness but weaker than the otherness found within the animal kingdom or the human one" (1990: 100). Ihde compares the "spirited sports car" and the "spirited horse". Both can be used as an instrument to travel, but the horse is capable of resistance and disobedience, whereas a car can only malfunction. A horse can also live on its own, while the car needs to be started to be "animated". Moreover, the car will not be startled when a rabbit suddenly crosses the road, just like most horses will not simply follow the "command" to run into a wall when the driver is too drunk to notice (Ihde, 1990: 99). The animation we ascribe to a car, like any other technology, is of a different sort than that of a horse (Verbeek, 2005: 127).

The last set of human-technology relation Ihde describes is the background relation. Here, technological artifacts do not play a central role in our experience or interactions with the world, but reside in the background of it (Ihde, 1990: 111). The heating and cooling by automated thermostats, the light of lamps and the zooming of our computers are all examples of technologies with which we form background relations. While their effects may be subtle and often go unnoticed, their role is to influence how we experience the world. Ihde schematises background relations as:

Background relations: Human (technology/World)

4.2 AUTHORITY RELATIONS WITH TECHNOLOGIES

Verbeek (2008; 2015) has already observed that many recent technologies do not fit into Ihde's categories. For instance, a brain implant that is used for deep brain stimulation to treat certain diseases or psychiatric disorders, is more than embodied; it merges with the human body into a new, hybrid being. Verbeek calls this a cyborg relation: human/technology → World (Verbeek, 2015: 29). Another example are wearable technologies such as smart glasses that can be embodied to give an experience of the world, while also giving a representation of the world in a parallel screen. Verbeek speaks of this combined embodiment and hermeneutic relation as *augmentation*:

(Human-technology) → World + Human → (technology – World).

Also, authority relations with technologies do not fit Ihde's human-technology relationships. On the one hand, technological authorities play a mediating role in the relation between humans and their world by actively shaping perceptions and actions. Such technologies tell people to think or do something by presenting them with a set of instructions or representations of reality ("your destination is on your right", "release inmate", "this is a risky individual", etc.), resembling a hermeneutic form of mediation that shapes experiences and interpretations which inform behavior. But with authority, on the other hand, the technology becomes the terminus of our experience in similar ways as with alterity relations. To recognise authority means to be directed at another actor that tells you what to think or do. This implies that not all technologies are potential authorities. A pair of glasses, a table or a broom will not trigger people to categorise them as authorities as they are not a source of interaction; but with digital technologies we do experience a sense of otherness based on their apparent autonomy and the interactions they give rise to. Yet, authority relations with technologies go beyond simple alterity as well. As discussed in chapter 3, it is the authoritative identity of an actor that motivates people to accept in advance whatever is communicated. In the authority relation, the influence that technologies exert on human perceptions and behaviors is so compelling that it becomes characteristic for the relationship between human beings, technologies, and their world.

As such, the digital authority relation is not so much about our interpretations and behaviors in the world *with* or *to* a technology but *by* or *according to* a technology. This could be schematically indicated as[65]:

Human → technology → human → World

4.3 WHAT THINGS SAY

It has already been stated several times that authority is considered an essentially social concept. But what makes this so? What is it about authority that would confine it to the realm of humans? It is in our use of language, or put more broadly, in the way we communicate. In paragraph 3.2, I discussed the work of Bruce Lincoln (1994) who linked authority relations to the capacity to produce consequential communications. Authority is an act of communication that incites accepting and compliant attitudes of an audience. And since such communications could always only be issued by people, authority remained tied to human relationships.

This is not to say that things did not say things. Scholars have also highlighted that many technologies send out messages on action possibilities. In paragraph 1.4, the notion of *script* was already introduced, referring to the prescriptions in technologies that steer users in how to relate, adapt and interpret a technology (Latour, 1992; Akrich, 1992). Another, well-known concept is that of *affordance*. Donald Norman (2013) describes an affordance as a design feature of an object that suggests how the object could possibly be used. Many of his examples concern buttons, doors and switches that are cognitively associated with possible actions. And in the case problems arise – people that push the wrong side of a door or when automatic doors open unexpectedly – it is not necessarily the user's fault but rather the technology that was designed wrong and conveyed the wrong signal, Norman says (2013). That is, an affordance that would direct the users towards suitable behavior apparently is missing. Furthermore, Thaler and Sunstein (2008) have put forward the concept of *nudge* for situations were technologies make subtle changes in the way choices and information are presented with the aim of guiding people towards desired choices and behaviors. Defaults are a good example of how individuals might be nudged by digital technologies into certain behaviors – think of tipping defaults and privacy settings.

Still, these are not the communications emblematic of authority. The messages of these technologies that guide and influence human behavior are implicit; they only communicate in a metaphorical sense. They are prescriptions encoded into the technology, inviting, or stimulating behaviors without making apparent claims. Notwithstanding the "affordances" of analogue technologies that amplify or reduce aspects of experience or the scripts that actively invite certain actions while discouraging others, such technologies do not explicitly command or advice (Van de Voort, Pieters and Consoli, 2015:42-4). For this reason, we should not speak of the authority of, let us say, hammers, canes, or coffee cups - they did not issue communications to be believed or followed. While the script of a speed bump can have a similar effect as the words of a police officer, we do not recognise it as an authority. The communications of a speed bump are implicit, and we tend not to focus on it too directly. And when subjected to nudges, we are often unaware of the influence that is exerted on us in the first place. With authority, however, the communications to be believed or followed are actively attributed to the other – in this case to the technology that communicates something and, as such, becomes a reason for submitting to it.

One question that arises here is whether we are really directed at the technologies in the first place. Do we recognise the communications coming from a technology, or is it their makers who we are hearing? Doing what a technology tells us can also highlight the authority of the designers using other means, rather than that of the technology. After all, technologies are designed and (still) lack *intentionality* (Searle, 1980) – no matter how intelligent they may act, they do not "know" what they are doing. While stop signs and traffic lights tells us to stop, we may decide to do so not because of the technology itself, but out of trust and respect for those who made them and the possible consequences for non-compliance. But the emergence of digital technologies potentially challenges this. This claim is not necessarily grounded in the extreme cases of Artificial Intelligence that make headlines, in which individuals have seen the "ghost in the machine" (The Washington Post, 2022). Such cases are about whether AI, now or in in the future, can develop consciousness. Surely, when technologies would, for example, indeed become conscious or truly emotionally intelligent – having or exhibiting innate emotional capabilities (see Picard, 1997) – one could argue that it is the thing itself that has become the one communicating as it has detached itself, like a child from its parents, from any human designer. Instead, building on the previous paragraph, I argue that we attribute all sorts of communications to less spectacular digital technologies, too. When interacting with many digital technologies, we tend to respond to their sounds, spoken words or written texts as if it were communications of their own – regardless of the human programmer lingering in the background. It becomes the technology that communicates to us. This is discussed in more detail in the next chapter.

Let us conclude, for now, that many human-technology relations are now characterised by digital communications that are actively responded to by human beings. I now continue the discussion on authority relations and its characteristics. In the previous chapter, the authority relationship was characterised as unequal, or fundamentally asymmetrical which produces attitudes of trust, respect, and acceptance. The rest of this chapter is dedicated to better understanding the potential nature and origins of this asymmetry when thinking about digital technologies. Let's start this off with a discussion on parking.

4.4 THE EVOLUTION OF PARKING ENFORCEMENT

Parking used to mean exactly what it sounds like: creating a park. In the spring of 1870, a law was passed by the United States Congress that authorised the city of Washington D.C. to reserve up to 50 percent of a street's width for the creation of "parks for trees and walks" (Richmond, 2015: 20). In the years that followed more than 70,0000 trees were planted on the side of the streets of Washington. The first parking system that followed thus had nothing to do with cars but was as Richmond observes "an early street tree system where parking defined the planning of trees, gras, and flowers along the side of roadways and the creation of sidewalks for pedestrians" (ibid.).

But it did not take long before the emerging roadside parks were used for more than just walking. People started tying up their horses and carriages to the parking trees when they were out in stores or visiting friends. This became the new "parking" and while it quickly became illegal to do so because it would hinder traffic and injure the trees, many took the risk of getting fined because of the convenience and shade that the trees offered for the waiting horses and carriages.

In the twentieth century cars came to dominate the streets. In the United States, the number of motor vehicles increased from approximately 8,000 in 1900 to more than 26 million in 1930 (FHWA, 2024). And just as people had tied their horses to a tree before, they now stopped their cars next to the parking strips on the side of the streets. City officials started cutting down trees and widening streets to deal with the sheer number of cars, thereby officially replacing the initial meaning of parking as a place where trees and bushes were planted with parking as a place to leave a car (Richmond, 2015: 21).

As car ownership increased, on-street parking quickly became scarce. Drivers could no longer park whenever and wherever they wanted, leading them to cruise around looking for a vacant space, wasting time, fuel, and congesting traffic (Shoup, 2011:1). Many municipalities struggled with the disordered bunch of parked cars and the city of Los Angeles even at some point experimented with banning on-street parking completely during daytime hours (Marshall, 2014: 362).[66] Other cities also tried to regulate parking, for example with posted parking time limits, but it turned out that these were very difficult to monitor. After all, officers could only observe and memorise the parking behavior of a limited number of vehicles within a certain area.

The introduction of a new technology in the 1930s, the parking meter, offered a solution as it allowed for keeping track of the time a vehicle was parked in a specific place. After inserting coins into the device, someone could park his or her car in the spot for a certain amount of time before more money must be inserted. When the person failed to do so before the time ran out, he or she could get a fine from observant officers checking the meters.[67]

Over time, most coin-operated, single-space, parking meters were replaced by electronic devices which covered many spaces within a defined area (Ison & Budd, 2016: 151). Initially, people needed to indicate on an electronic screen for how long they were planning to use the space, pay by cash or card and then received a printed ticket to be displayed in their vehicle. With newer systems, however, a printed ticket is no longer required. Modern parking machines and parking applications on mobile phones only require drivers to enter their vehicles' license plate. Parking officers, in turn, walk the streets and use handheld computers to check the plates of the vehicles parked in a particular area to ensure the parking session is valid.

Meanwhile, the next step in the evolution of parking has announced itself: scan cars. Scan cars are equipped with cameras on the roof that, while driving, read the license plates of parked vehicles to detect parking violations. The system compares the scanned plates to a database of paid parking transactions. Subsequently, in the case of complete digital enforcement, fines are sent automatically to the homes of the owners of the vehicles that are not found in the database. In many cities, however, human officers are (still) the ones deciding about the follow-up. The technology's alert of an illegally parked vehicle is then sent to a desk-officer who checks the scans and decides to issue a parking ticket or to dispatch a parking enforcement officer on a scooter to evaluate the situation on the street – for example to make sure that someone was not buying a ticket at the time the car was being scanned or to check for emergencies and other exceptional circumstances.[68] When the officer, in turn, determines that a car is, in fact, parked illegally, a fine will be issued via the system and sent to the registered owner.

A scan car reads more than 1,000 license plates per hour. This has not only lowered enforcement costs but also tremendously boosted city's incomes: digitisation resulted in an increase in parking fines. Moreover, the use of scan cars influenced motorists' behavior: as the chances of being caught increased, more and more motorists decided to pay for their parking or stopped parking in paid areas. It is therefore no surprise more and more cities throughout the world are adopting scan cars for parking enforcement.

Technological innovation has been central in the evolution of parking as well in the work of those enforcing parking regulations. When parking spaces became a scarcer commodity and regulation became a necessary evil, officers became responsible for ensuring that drivers would comply with local parking laws and respond to violations. This initially meant relying on human senses and memory. Later, parking meters supplemented officers' observations. Coin-operated meters and printed tickets placed behind vehicles' windscreens communicated (non-)compliance, allowing for more effective modes of patrolling. Consequently, parking enforcement officers became constrained in acquiring information themselves as they now covered larger areas, making it increasingly difficult for them to monitor every car. With the introduction of digital systems where drivers register their license plate and sign up for a virtual parking ticket, information resources available to officers got limited further. For determining parking (non-)compliance, enforcement officers now needed to rely on that what is presented to them on the screen of their handheld computer. While they might, coincidentally, have seen someone buy a ticket themselves allowing them to assess the adequacy of what is in the system, most of the time they are unable to do so as they are patrolling large areas and physical evidence (e.g., a printed ticket or a meter on the sidewalk) has been designed away. With the introduction of scan cars, the relationship between enforcement officers and the availability of information in their work-environment has become even further obscured. Officers are no longer found wandering the streets but are only "on the spot" after they have been alerted by the digital technology. Their work is now highly structured around the communications of a digital technology. They no longer have a general overview of parking in the city themselves, but only through their computer. For the modern parking officer, additional information can only be acquired after they arrive at the place where the system has sent them.

4.5 THE ONES TO SPEAK IN AN EPISTEMIC NICHE

The story of parking makes visible how the relationship with digital technologies has become increasingly central in people's interactions with the world. With the introduction of digital technologies, we frequently see how the structure and sources of knowledge available to human decision-makers get severely constrained. Philosopher Jeroen van der Hoven (1998:100) speaks of *epistemic niches*. Epistemic niches are characterised by a structural knowledge-scarcity, making it difficult to get around an information system: users must decide, often in short timeframes and are constrained in acquiring relevant insights from outside the niche. An air traffic controller in a tower, for example, has as his most important source of information that what is presented to him on a screen (e.g., flight paths, altitudes, airspeeds, etc.). Similarly, modern-day parking enforcement officers are primarily informed about vehicles' valid parking tickets by a handheld computer. For them the possibility to make observations themselves and to acquire useful additional information is severely limited. In the case of the air traffic controller human senses just do not allow for it in the first place - they generally cannot see where planes are flying, at what altitude and how fast they are going. Digital technologies have become indispensable for their actions and decisions.[69] In other situations, relevant information can be acquired elsewhere, but the structure and nature of people's work may prevent them from doing so. While users of digital technologies can occasionally call and check with a colleague or do an algorithms' calculations manually, – as was done in Washington State prisons after the software mistake responsible for the wrongful release of thousands of inmates was discovered - such strategies do not seem to be feasible in the long run as they take up a lot of time and are not flawless either. Organisational rules and structures can make alternative methods of obtaining information not a very attractive alternative. Leaving a digital technology as pretty much the only ones to turn too.

Epistemic niches tend to promote relations of *epistemic dependence*. This notion was developed by John Hardwig (1985) and refers to the situation where people come to rely on the knowledge of others but cannot assess its truth themselves. For them the belief that someone (or something) else has good reasons to believe something is true, forms, in itself, a good reason to believe in it.[70] For example, we depend on physicists with respect to the truth of Einstein's theory of relativity or on medical specialists for the truth of their diagnoses (Hardwig, 1985: 339). Most of us lack the training and ability that is required to assess whether relativity theory is correct or that allows us to make our own medical diagnoses. Rather, we believe that Einstein or a doctor has good reasons for believing in their claims, and therefore believe in their claims too. In a similar way, users can come to epistemically depend on a digital technology: they cannot assess the truth of what the digital system they are working with claims to be the case. People using digital technologies face the problem of opacity and often cannot check or verify the system's communications (Hayes et al., 2020: 3). Very often users have no idea based on what an algorithm advises them. Still, they might have good reasons to accept what the system says – for instance because it is known to be reliable or when there is simply no other information available to them.[71]

In short, the use of digital technologies may come with an artificial knowledge vacuum: an epistemic niche. Here, the system presents its output as accurate and worthy of belief, while at the same time the user becomes constrained in acquiring knowledge from alternative sources (Van den Hoven, 1998: 103). With limited alternative sources of information available to those in an epistemic niche, individuals may come to depend on a digital technology, with little or no control over the accuracy of the information provided. They become epistemically dependent.[72] The outcome is an asymmetry between the technology and the individual; with limited alternative sources of information available to us, our relationship to digital technologies comes to the front of our experience. Here, computers seem to have become pretty much the only ones to speak. In an epistemic niche, the way a digital technology presents the world, will be pretty much how the world will come to appear to its audience. As such, when decisions must be made and a technology convincingly presents its output, it becomes very likely that people will follow it.[73] It might even be tempting to say that in such environments digital technologies have authority by default – but this would be premature conclusion as I show in chapter 7.[74]

4.6 PERSUASION IN TECHNOLOGY

Oftentimes, however, there is no pure epistemic niche shaping the relationship between humans and digital technologies.[75] The rivers and lakes in which people drive their cars are visible to everyone. And a security officer can deduce that a 2-year-old is probably not a terrorist (see next chapter). What, then, can we find, in the human-technology relation, that would explain people's reliance on digital communications rather than on their own intuitions or the ideas of others? The remainder of this chapter elaborates more on the possibilities of persuasion and coercion in human-technology relations as constitutive of the asymmetry characteristic of authority relations. I start with the former.

Persuasion is an age-old phenomenon that in the last decades has also become linked to technology. In his book, *Persuasive Technology* (2003), Fogg suggests that technological artifacts can persuade people, just like discursive arguments, to change their attitudes and behavior. To change what people think and do, technologies can, for instance, reward or remind them of (un)desirable behaviors. Think of commercial websites that provide tailored suggestions about books you might want to buy or the screen next to the road that shows a sad face when you're driving too fast. Fogg also mentions a system which monitors handwashing in restrooms and sends a signal in case people forget (Fogg, 2003: 47). With the development of this strand of research digital technologies are now increasingly recognised as potentially persuasive actors. In fact, Fogg writes that computers have several advantages over human persuaders (Fogg, 2003: 7). Among other things, computers can be more persistent and do not get tired, they can draw on a vast storehouse of information and they may be ubiquitous – having the ability to be almost everywhere (ibid, 7-11). According to Fogg, these characteristics undoubtedly increase the persuasive powers of a technology.

But authorities do not need to persuade. Persuasion only is needed when there is no authority. Yet, the possibility of persuasion, when present in the background, is fundamental for constructing such relationships. As argued in chapter 3, the exercise of authority may hinge on the presumption that persuasive argument can and will be given when asked for. The work on persuasive technology, and our own encounters with such technologies, highlights how persuasion can be considered as potentiality within human-technology relations. Especially the general ability of computers to access and manipulate large volumes of information is of significance here. People often presume that there is a world of data and complex algorithms underlying the digital output that is presented to them and that precisely the right fact, statistic, or reference from that volume of data can be brought forward to persuade (Fogg, 2003: 8-9). The computer is a *black box* to them. Users themselves may or may not be barred from these insights themselves. But it is the belief in the existence of good reasons for compliance and that these can be provided when asked to do so – a capacity for reasoned elaboration – that underlies people's compliant responses. The inequality that underlies digital authority, then, is grounded in the perception that one (the technology) is has superior access to knowledge than the subject itself – the experience that "some are wise and some are otherwise".[76]

Still, the actual persuasive powers of digital technologies can be rather static. Often a computer will not provide users with exact explanations for their output because it is not programmed to respond as such. When people want an explanation, they generally must search for one themselves – for example by looking into the data or investigating the algorithm. Or people may not understand the digital explanation that is given – bringing us back to the idea of epistemic dependence as discussed above. Moreover, the elaboration that a digital technology provides us with often requires much interpretation. The beep, blinking light, or brief written text a technology might present us, often resonates because it triggers elaborative schemata already present within us – it incites us to tell ourselves that it is safer to wear seatbelts, that there is a world of evidence suggesting that taking regular breaks from typing is good for your health, that the system is right most of the time, etc.

And very often, the technological capacity to persuade is mediated – depending on the presence of others. Frequently it is other people that would provide us with reasons why we should do or think the thing the technology says. Programmers, designers, system managers, implementation officers and the like tend to function as henchmen of digital authority and are thought to be capable of providing further explanation upon request. Again, not everyone will understand them, but it is more about the *belief* that logical evidence for digital claims exists somewhere. This idea can motivate people to follow the technology regardless. When users assume that good reasons exist for whatever they are told, they end up being persuaded in advance. As Lincoln concluded, authority, then, becomes a "time-saving device or shorthand version of persuasion" (Lincoln, 1994: 5).

4.7 AUTHORISATION AND COERCION IN TECHNOLOGY

Digital technologies do not hold official positions with formalised powers. Unlike police officers, schoolteachers, managers, or military officials they do not have formal capacities that could enforce behavior. Computers cannot fire, arrest, or ban people. In fact, the General Data Protection Regulation (GDPR) does not allow for automated individual decision making at all. As stated in article 22 (1):

> "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her".

So far, there is no law that grants digital technologies the legal right to rule.[77] However, institutions do involve them in many decision-making processes and grant them positions of influence. As discussed in more detail in the next chapter, digital technologies are now advanced to various formal roles that are typically associated with authority. In these roles, they are integrated into institutional networks that foster trust and respect for those who are part of them. The asymmetry between the technology and human subjects, then, is institutionally ascribed and the result of a process of *authorisation*. Here, digital technologies have gained the confidence or support of an institution or its relevant members (such as managers or leaders) who confer authority upon the technology. This can take similar forms as with human authority. For example, a store manager can decide to hire a store security guard to prevent theft, or a coach can select a team captain for leadership on the field. When authorised by respected others, they are able to attract attention and evoke compliance from customers or players. Similarly, an organisation can confer authority on an expert system by making it mandatory for staff to consult it when making decisions. While someone might still want to refute the output of the technology, a narrative of institutional confidence can cultivate respectful and trusting attitudes towards the technology, highlighting its special position, and ultimately discouraging opposition.

Still, this perspective already highlights the dependence of digital authority on humans. While there have been some phenomenological explorations of human submission to digital leaders – specifically robots (Gladden, 2014) - people still hold positions that enable them to promote and restrict access to position of authority. This is also true when thinking about the potentiality for coercion in human-technology relationships. Here too, coercion is often conditional and depends again on the support of humans in positions of power. Only through the active or passive involvement of another authority figure, the potentiality of force becomes present in the background of the human-technology relation. Individuals interacting with the technology can come to presume that their resistance will be logged somewhere and that ignoring the digital output could get them into trouble.

But down the line, the capacity of coercive force rests with the backing of a human authority figure and when they withdraw support, the technologies' access to coercive force fades.

However, there also appear to be digital technologies that can exert force against individuals who refuse to comply with instructions or demands more autonomously. This takes us to the domain of *techno-regulation*. In his famous book *Code 2.0*, Lawrence Lessig (2006) observes how legal scholars tend to use legal rules to remedy pressing societal problems but that there are also many other regulatory forces that may shape human choices and behaviors such as market forces (e.g., taxation or price mechanisms) or social norms. In similar ways, Lessig argues, individuals' behaviors are shaped by architectures (Lessig, 2006: 123). Not only are we affected by architecture in the literal sense of the word – the buildings and structures that physically organise our environment and as such affect what we can and cannot do -, we are also influenced by architecture in a more figurative sense. Lessig says that software and hardware are forms of architecture that "constitute a set of constraints on how you can behave" (Lessig, 2006: 124). As put elsewhere, they are forces that steer, guide, and influence behaviors and actions of end users (Van den Berg & Keymolen, 2017: 191).

The idea that rules and norms can be implemented into technologies to regulate people's behaviors is known as *techno-regulation* (Brownsword, 2004; Koops, 2011; Hildebrandt, 2011; Leenes, 2011; Van den Berg & Leenes, 2013; Van den Berg & Keymolen, 2017). The design of software and hardware can be used to deliberately influence the actions that are available to users. They can only use the technology in certain ways, predetermined by the designers or regulators. Any option of non-confirming behavior is designed out - maybe except for those higher up in the hierarchy - or non-compliance is made utterly unattractive. In essence, techno-regulation binds users through norms that are embedded in the technology thereby aiming to prevent or discourage undesired or illegal behavior. This also makes that techno-regulation is effective; individuals may attempt to resist compliance, but they will ultimately be compelled to adhere regardless.

An often-used example of techno-regulation is that of speed bumps (Latour, 1992; Brownsword & Yeung, 2008; Leenes, 2011). While there are different means to regulate traffic speeds, such as traffic signs, speed bumps are effective because they do so in a direct and binding way: they leave little room for drivers to divert from a regulator's intended course of action (Van den Berg, 2011:405). Surely, a driver may choose to be "disobedient" and drive over the speed bump at a high speed, but this is uncomfortable and may damage the car – unlike the traffic sign which can more easily be ignored. As a result, with speed bumps drivers are more likely to adhere to the traffic rules than when a traffic sign is posted (ibid.).

In a similar vein, techno-regulation has become closely associated with digital technologies. The computerised access gate can be seen as a digital equivalent of the speed bump. The gate only opens for people who are allowed to access, such as employees, individuals with valid passports or tickets or those with appointments. For those not in the system, the gate remains closed. Some individuals might, in turn, try to gain access by climbing over or forcing their way through the gate, but this can be dangerous, painful, or physically impossible. Like a speed bump, a computerised gate effectively regulates the behavior of those encountering it.

Techno-regulation offers interesting insights into the potential for coercion in digital technologies. These powers are vested in the digital technology or system itself; no others are needed to execute them (Leenes, 2011: 147).[78] Humans are only needed in the initial stages to design and authorise the technology to regulates human behavior in specific ways. It reveals how some technological artifacts can secure compliance with the alternative of force present in the background of the interaction. Even when an individual disagrees, they may still choose to follow the instructions of the technology because they understand that failure to comply will result in coercion – the gate will not open, the form will not save, or the system would simply not respond. This potential for coercive force can lead individuals to unquestionably do as they are digitally told.[79] Especially with technology, there tends to be no room for negotiation or bargaining. The technology can and will use force if it has too - resistance will be met with similar responses every time. This also makes the possibility of coercion a potential element for the experienced asymmetry on which digital authority is built, securing the digital capacity to produce consequential communications. This, however, does not mean that digital authorities structurally resort to the use of force. Both human and digital authorities do not need to persuade or coerce. The moment it becomes necessary to rely on coercion or persuasion to get something done there is no longer authority but just influence through force or arguments. To be a constructive feature of digital authority, the potentiality of coercion embedded into digital technologies through techno-regulation must remain exactly that: a possibility latently present in the interaction between the technology and the user. The possibility that a technology will use force when confronted with disobedience may contribute to an individual's regard for the technology and the willingness to accept whatever it communicates.