

Man as 'aggregate of data'

Sjoukje van der Meulen & Max Bruinsma

AI & SOCIETY

Journal of Knowledge, Culture and
Communication

ISSN 0951-5666

AI & Soc

DOI 10.1007/s00146-018-0852-6



Volu

AI & SOCIETY

Knowledge, Culture and Communication

25th Anniversary Volume

A Faustian Exchange: What is it to be human in the era of Ubiquitous Technology?

Karamjit S. Gill, *Editor*

Victoria Vesna, *North American Editor*

David Smith, *Open Forum Editor*

Richard Ennals, *Review Editor*

Mike Cooley, *Chair, Advisory Board*

 Springer

 Springer

Your article is published under the Creative Commons Attribution license which allows users to read, copy, distribute and make derivative works, as long as the author of the original work is cited. You may self-archive this article on your own website, an institutional repository or funder's repository and make it publicly available immediately.



Man as ‘aggregate of data’

What computers shouldn’t do

Sjoukje van der Meulen¹ · Max Bruinsma²

Received: 4 October 2017 / Accepted: 10 June 2018
© The Author(s) 2018

Abstract

Since the emergence of the innovative field of artificial intelligence (AI) in the 1960s, the late Hubert Dreyfus insisted on the ontological distinction between man and machine, human and artificial intelligence. In the different editions of his classic and influential book *What computers can’t do* (1972), he posits that an algorithmic machine can never fully simulate the complex functioning of the human mind—not now, nor in the future. Dreyfus’ categorical distinctions between man and machine are still relevant today, but their relation has become more complex in our increasingly data-driven society. We, humans, are continuously immersed within a technological universe, while at the same time ubiquitous computing, in the words of computer scientist Mark Weiser, “forces computers to live out here in the world with people” (De Souza e Silva in *Interfaces of hybrid spaces*. In: Kavoori AP, Arceneaux N (eds) *The cell phone reader*. Peter Lang Publishing, New York, 2006, p 20). Dreyfus’ ideas are therefore challenged by thinkers such as Weiser, Kevin Kelly, Bruno Latour, Philip Agre, and Peter Paul Verbeek, who all argue that humans are much more intrinsically linked to machines than the original dichotomy suggests—they have evolved in concert. Through a discussion of the classical concepts of *individuum* and ‘authenticity’ within Western civilization, this paper argues that within the ever-expanding data-sphere of the twenty-first century, a new concept of man as ‘aggregate of data’ has emerged, which further erodes and undermines the categorical distinction between man and machine. This raises political and ethical questions beyond the confines of technology and artificial intelligence. Moreover, this seemingly never-ending debate on what computers should (or should not) do provokes the philosophical necessity to once again define the concept of what it is to be ‘human.’

Keywords Hubert Dreyfus · Artificial intelligence (AI) · Man as ‘aggregate of data’ · Individuality · Authenticity · Philosophy and ethics

1 Introduction

Already in the first pages of the first edition of *What Computers can’t do* (1972), Hubert Dreyfus connects what he describes as the philosophical origin of AI to morality.¹ “The story of artificial intelligence,” Dreyfus writes, “might well

begin around 450 BC when (according to Plato) Socrates demands of Euthyphro, a fellow Athenian who, in the name of piety, is about to turn in his own father for murder: ‘I want to know what is characteristic of piety which makes all actions pious... that I may have it to turn to, and to use as a standard whereby to judge your actions and those of other man’” (Dreyfus 1993, p. 67). And Dreyfus continues: “Socrates is asking Euthyphro for what modern computer theorists would call an ‘effective procedure,’ a set of rules which tells us, from moment to moment, precisely how to

✉ Sjoukje van der Meulen
s.vandermeulen@uu.nl

Max Bruinsma
maxb@maxbruinsma.nl

¹ Department of History and Art History, Utrecht University, Drift 6, 3512 BS Utrecht, The Netherlands

² Camera Arts, Lucerne School of Art and Design, Sentimatt 1, Dammstrasse, 6003 Lucerne, Switzerland

¹ The first edition of *What computers can’t do. A critique of artificial reason* was published in 1972 by Harper & Row Publishers. The edition used in this paper is the MIT Press Edition from 1993 (second printing of the 1992 edition), with the slightly revised title *What computers still can’t do* and a new introduction by Dreyfus himself. An earlier revised edition of the book came out in 1979.

behave” (p. 67). The rest of the book, as we know, is a critique of exactly this optimistic idea—common in AI research circles of the 1960s and early 1970s—that everything concerning the operations of the human mind and behavior can be formalized as such an ‘effective procedure.’ In Dreyfus’ view, of course, biological and socio-political man is a much more complex, and, above all, embodied being. In *Alchemy and Artificial Intelligence* (1965), Dreyfus criticized leading artificial intelligence researchers for not fully grasping the “little understood human mind.” He went on to develop a systematic critique in *What Computers Can’t do* (1972), in which he argued against common assumptions in AI research holding that man’s cognitive processes could be emulated in algorithmic code, and that computers could thus be turned into intelligent machines. In the same book, Dreyfus also presented his alternative view on the human–machine relation, highlighting the role of the physical body in our experiencing and understanding the world, with which we are interacting as embodied species. In the decades after these key publications, Dreyfus deepened his philosophical position in regard to the man–machine dichotomy through his continued study of the work of French phenomenologist Maurice Merleau-Ponty and German philosopher Martin Heidegger, in books such as *Mind over Machine* (1986) and *Being in the World* (1991). Towards the end of his career, Dreyfus extended his technological and philosophical insights to the newest computational domain of disembodied presence: the Internet. Dreyfus’ *On The Internet* (2001) is one of the first sustained critiques of this new medium, exploring the virtues and limits of the (early) World Wide Web from the perspective of internet-intrinsic characteristics such as hyperlinks, interactive virtual environments, and even what he considers to be the false premises of online learning, due to the latter’s flawed ambition to completely replace the embodied classroom.² The root cause of the error that most AI technophiles make, according to Dreyfus, is a lack of understanding of the fundamentally different nature of man and machine, which leads to serious moral issues concerning humanity, and is thus the reason why he insisted until his death in April 2017 that we should reflect on the concept of man more deeply and philosophically in the current age.

‘Man’ is a conceptual composite from the very start of ancient Greek philosophy, which spawned what can broadly be summarized as the body–mind problem. Beyond this basic dualism between the physical body and the immaterial mind (or ‘soul’ in Greek philosophy), Plato already further subdivided the soul into three parts to come to terms with

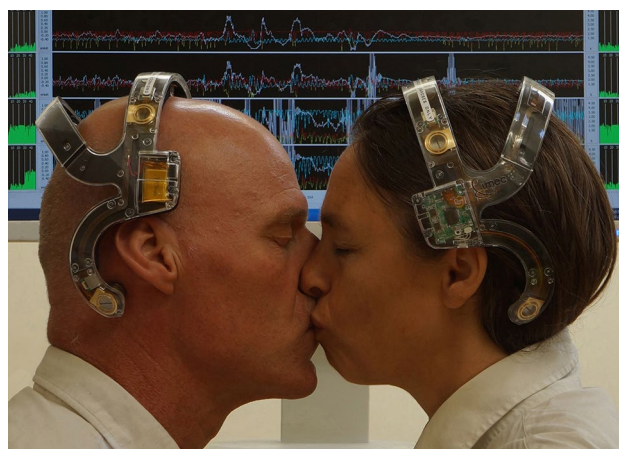


Fig. 1 E.E.G. KISS. Credits: © Karen Lancel & Hermen Maat (LancelMaat). 2014. <http://www.lancelmaat.nl/work/e.e.g-kiss/>

the complexities of the human psyche: a rational part (the mind), a spirited part (emotions related to sense of justice) and an appetitive part (which satisfies bodily drives and needs). The great Christian theologian Augustine redefined the relationship between body and soul by adding an entirely new third concept—the flesh—to basically suit the doctrine of sin. Or in Augustine’s unforgettable words: “And it was not the corruptible flesh that made the soul sinful; it was the sinful soul that made the flesh corruptible” (Augustine, p. 551). In a more scientific mind, René Descartes drew a clear-cut line between the physical body and the immaterial mind, and from then on we identify the mind–body problem with the notion of ‘dualism’, which produced yet another pair of terms: ‘mind’ (immaterial) versus ‘brain’ (physical). And so on. The philosophical debate on the mind–body dichotomy is long and complex, but what is at stake in this essay on the legacy of Dreyfus’ work is the claim that something essential has changed about our concept of man in the information age: a new type or ‘part’ of man, which functions and acts at the nexus of his physical and immaterial manifestations—deeply affecting both—needs to be taken into consideration: ‘Man as aggregate of data’ (Fig. 1).

2 From individual to dataviduals

Imagine someone with a home, a job, friends and a number of habits and opinions. They could move to another house, change their habits, start doing other work, boost their social life, and nudge their mind into another direction. Our imaginary person could change sex and alter their appearance, if they wanted, or disappear into the crowd anonymously, faceless and genderless. If someone could do all of those things, who is he or she now? A free individual or a packet

² The first edition came out in 2001, a second, revised version was published in 2011 (including for example an additional chapter on the online 3D virtual world, ‘Second Life’).

of free-floating data? Analyzing the concept of what it is to be human today, we are all but forced to conclude the latter. On the surface, it seems as if we are living in hyper-individualistic times, but on closer look our concept of ‘individual’ appears to be about ‘unicity’ in a rather different way today than this term traditionally suggested. The ‘indivisibility’—the transliteration of the Latin word *individuum*—of the individual is traditionally associated with a unique personality. But, if we look at today’s concept of man, the individual is more multi-part part of, than undivided. So, what does that mean for ‘me’ as ‘unique’ person?

At the dawn of personal computing and the Internet in the early 1990s, the philosopher Gilles Deleuze envisioned the transition from the notion of ‘individual’ to ‘dividual,’ within the broad societal change from modern ‘disciplinary society’ (as defined by Michel Foucault) to what Deleuze coins ‘the society of control’:

The disciplinary societies have two poles: the signature that designates the individual, and the number or administrative numeration that indicates his or her position within a mass. This is because the disciplines never saw any incompatibility between these two, and because at the same time power individualizes and masses together, that is, constitutes those over whom it exercises power into a body and molds the individuality of each member of that body. (...) In the societies of control, on the other hand, what is important is no longer either a signature or a number, but a code: the code is a password, while on the other hand disciplinary societies are regulated by watchwords (...). The numerical language of control is made of codes that mark access to information, or reject it. We no longer find ourselves dealing with the mass/individual pair. Individuals have become “dividuals,” and masses, samples, data, markets, or “banks” (Deleuze 1992, p. 5).

In today’s information- and network society, we argue in this paper, the individual is not only part of (and particle within) these new overarching structures of control, he is also compiled of (data) components which characterize him and which are in a sense interchangeable for other components. These are his data, information of all sorts, which together result in a more or less unique profile. The unicity of today’s ‘I,’ therefore, is due to the improbability of there being another combination of data, exactly like ‘mine.’

An intermediate conclusion of this description of a concept of man built on data is that if one looks at it this way, everything can be different as well. Even if some data will stay the same, one’s profile changes a lot when one starts eating, reading, and buying differently. In other words: how constant is the individual? Again, from the vantage point of data analysis, or more accurately, statistics, one can see patterns: similar things one has been doing or using for years. Our imaginary individual might, for instance, have used only three mobile phones in 20 years, or liked vegetarian food already for a very long time. It seems trivial, but with the

right algorithms one can infer a lot from such data—or at least many presume they can—mainly by looking at how this seemingly unique profile corresponds to someone else’s. It hammers home that individuals are constantly being looked at as belonging to a category. Regardless of all rhetoric about how *you* rule as individual, you are constantly sliced up in sub-categories, each the playing field of institutions (governments, companies, interest groups) who see, trace and monitor you as target. Despite your individuality, you are primarily seen as part of a target group, which in principle consists of a well-described and confined stock of ‘dataviduals’.

Of course, this has in some sense always been the case. Man is a social being, and therefore entwined with greater frameworks. He is part of a group. Still, the seemingly subtle difference between ‘group’ and ‘target group’ tells a lot about how our society has changed in the modern period—Theodor Adorno and Max Horkheimer already complained about this in their famous chapter on the ‘Culture Industry’ (Adorno and Horkheimer 2002)³, and the digital age has sharpened this distinction even more. The inevitability of being oneself within the group into which one was born has given way to the idea that one chooses the group one wants to belong to in both the physical and digital world. This choice, and most of all the social pressure to choose—for not belonging to a group is as socially undesirable as it always has been—has led to an intensified searching for means to express this identification; not so much with oneself, but with the group of one’s choice. Thus, the group (within which one identifies oneself) becomes a target group (with which one wants to identify from the outside). And, there grows an entire industry in late-capitalist society to assist you, both online and offline, offering a range of products and services which allow you to identify with the group or lifestyle you aspire to. Which help you realize your personal identity, your authentic self. This, in turn, begs the question: how authentic are you, actually, if your sense of individuality is governed by the algorithms of the *Kulturindustrie*, rather than personal reflection and introspection?

3 The do-you-yourselfer

One of the vital roots of the Western concept of authenticity is the idea of personal freedom of choice—an authentic human being makes his own choices. Free choice already is

³ This book was originally published in German during the Second World War under the prosaic title *Philosophische Fragmente* by Social Studies Association, Inc., New York (1944). A revised edition was published in 1947 by Querido Publishers in Amsterdam under the current title *Dialektik der Aufklärung*. The authors’s criticism and warnings of the ways in which individuality is mass-produced and manipulated in our technologically driven, industrial culture is still pertinent for digital culture today.

a central issue in one of the oldest myths of Jewish Christian culture: the expulsion from Paradise. The first humans violated God's ban on tasting the fruit of the 'Tree of Knowledge of Good and Evil.' The gist of this story is that as a result man did not only learn to discern good from evil but at the same time became burdened with the task to choose between the two by his own free will. For the knowledge of the difference implies that you reflect on which way to go: for the good or the bad. Man is since assisted in this choice by countless advisors, in whom one often recognizes the figure of the snake from the original myth; the creature that, with his persuasive words and irresistible logic, seduced Eve to taste the doomed fruit, who in turn seduced Adam to join. From Augustine's defense of free will in man's combat of evil in *On the free choice of the will* (426 AD) to Martin Luther's denial of the free will in *The Bondage of the Will* (1525) in the reformation, the paradox of predestination versus free will has haunted religion—and Western culture in general (Augustine 2010; Luther 1957).

For free will is not only a curse in religion but also in philosophy, as is shown by the centuries old discourse of man as individual, as authentic and autonomous being, who chooses and upholds his own place in the world and among his fellow men. It resounds in Sartre's dictum '*l'Enfer, c'est les autres*' ("Hell is the others," Sartre 1945), implying that it is always 'the other' who thwarts and undermines our self-image, who forces us to face our own insufficiency. This human condition—the general condition that 'I' is always 'someone else' in the eyes of the others, who see themselves as 'I' as well, from which mathematically follows that all men minus one are wrong—makes us susceptible for seduction and manipulation, but also for collaboration. For when we work together, we coincide with the others and become a collective 'I.' Working together is contrasted by 'doing it yourself,' which in turn is a transliteration of the old Greek word *αυθέντης*, (authentēs, 'self-doing'), which is at the base of our word 'authentic.' In many variants since, say Socrates, this 'self-doing' is thought as a combination of reflection and action, which forms the core of each individual's personal responsibility, be it towards God or society—his conscience. Someone who thoughtlessly does what others instruct him to do is not 'authentic;' his actions are merely going along with the flow. He does what others tell or suggest him to do and doesn't reflect on whether he does right or wrong by that. The authentic man, however, the 'do-you-yourself,' decides for himself what he does, and why.

So far, so good. But on what does an authentic individual base his choices? This brings us back to the theological–philosophical Tree of Knowledge of Good and Evil. For even if, since Nietzsche, God is dead and does not supervise our dealings anymore, the difference between good and evil still exists, along with human constructions of ever-changing ideas of morality. And so does the autonomous individual's

responsibility to check his ethics and moral choices against it. The modern concept of authenticity dates from the eighteenth century, when it developed alongside individualism, in particular the notion that the human being is an autonomous agent directing his own will, as Kant described in his foundational ethics, *Groundwork of the Metaphysics of Morals* (Nietzsche 1967).⁴ Since Kant, many references for this free choice are related to his basic philosophical idea of *allgemein subjektive* (generally subjective) truths; truths that cannot be proven objectively, but which are still experienced as being categorically true and therefore morally operational. In Kant's ethical theory, such truths are more specifically defined as 'categorical imperative,' the kind of truths of which everyone intuitively knows which side they represent, good or bad (like murder, theft or lying), but which still need to be tested against autonomously determined principles or so-called 'maxims' on which a person acts, and which depend on situation and context. In Kant's words: "Act only in accordance with that maxim through which you can at the same time will that it become a universal law" (Kant 1998, p. 31).

One complication is that no one makes these grand universal truths authentically themselves. They are collective agreements, not individual findings. They are there, passed along for generations and sometimes, suddenly or gradually, appear to be subject to erosion. This erosion of truths that were thought of as permanent, by contrast, is often the work of 'authentic characters' in the Nietzschean sense of ascetic individuals who are, somehow, beyond established truths and morality—quite literally, *Jenseits von Gut und Böse* (Nietzsche 1886). People who critically reflect on their choices against the background of what 'the group,' society, deems morally desirable or decent. Often against the grain, but always in discussion, responding to the pressures of the material world and the others. In other words: an authentic person makes himself—with the means provided to him by his society, culture and belief.

The Western concept of authenticity is strongly connected to the old Greek advice *Γνώθι σαυτόν*, (Gnothi se auton, Know thyself). In antiquity, this not only meant that one should know and analyze one's own motives, emotions and actions, but also that this scrutiny was linked to what one's culture saw as its highest values. The self-reflection of the 'be-and-do-you-yourself' human being, the 'authentic' human, took place unambiguously within the cultural, political and spiritual institutions of the society of which one was an individual, undivided, part; an 'atom,' the smallest particle of a larger whole (the Greek word *ατομος* also means 'undivided,' just like the Latin *individuum*). This short excursion into the Greek and Latin roots of the modern

⁴ Originally published in 1785.

concept of ‘individual’—appropriate in a special honorary issue on Hubert Dreyfus, who first insisted that AI should consult philosophy from the classics (Aristotle and Plato) to the moderns (Heidegger, Merleau-Ponty, the late Wittgenstein) on all human matters (Dreyfus 1993, pp. 67–68, 252–253, 261–263)—indicates that our culture has, for a very long time already, understood the individual human person as the smallest common denominator of society, and not as a loose, unconnected particle.

4 The quantified self

This brings us back to the concept of man as aggregate of data, which we described above. The nagging question is whether, in today’s network society, one can still conceive of an authentic self in the classical sense of ‘knowing thyself.’ How can this new type of individual—or ‘dividual’ in Deleuze’s terms—that is primarily known (and knows himself) as aggregate of data and functions, still be called authentic? And what happens to the autonomous being, who does not wish to subordinate himself to what others, including computers and robots, expect him to do? In other words: is a hybrid human-data aggregate still a self-aware individual with the capacity of taking responsibility for his own thoughts and actions or are his mind and behavior more and more conditioned and controlled by algorithms and software codes, as Lev Manovich would have it:

I think of software as a layer that permeates all areas of contemporary societies. Therefore, if we want to understand contemporary techniques of control, communication, representation, simulation, analysis, decision making, memory, vision, writing, and interaction, our analysis can’t be complete until we consider this software layer (Manovich 2008, p. 7).⁵

Kevin Kelly, co-founder and former editor of Wired magazine, is the acknowledged specialist of data-assisted man and his power to acquire self-knowledge and self-awareness through new technologies, especially digital tracking tools that promise us to help improve our life, health and happiness. The underlying assumption is that new data-gathering software for computers, mobile phones and sensor based gadgets can change our sense of ‘Self’ and of ‘being in the world’.

Since data technologies have become tinier, cheaper and distributed on an unprecedented scale, it has indeed become easier to appropriate the quantitative methods first used in science and business and apply them to the social and personal sphere for health and wellness improvement,

⁵ http://softwarestudies.com/softbook/manovich_softbook_11_20_2008.pdf (accessed 15 July 2017).

tracking weight, physical activity, sleep patterns, professional productivity, etc. Together with partner Gary Wolf, Kelly came up with the concept of ‘Quantified Self’ at the first eponymous conference in San Francisco in 2007. The QS movement that developed out of this conference, has full confidence that systematic self-tracking will lead to data-assisted self-awareness and personal growth; or, as the dictum on the accompanying website says ‘self-knowledge through numbers.’⁶

Remarkably, the idea of the Quantified Self hinges on data (immaterial information, that is) generated by our physical bodies interacting with our techno-material environment. Far from forgetting that our being in the world is still very corporeal, the Quantified Self integrates our ‘flesh’ within what Stephen Humphreys and others have called ‘datasphere’,⁷ which comprises every aspect of our lives, both private and public. In the datasphere, our bodies virtually dissolve with any other representation of our existence, physical or abstract. This is not to say that we are becoming less carnal, but that the old dichotomy of body and mind is being radically redefined—bodies too become amendable in a conceptually different way than the ancient *mens sana in corpore sanum* suggested. The concept of man as aggregate of data, in short, does not make him less physical—it shifts the focus from the “flesh” to the ways in which our bodily movements and actions, including our most private ones, “take on materiality as they become artefacts of the datasphere” (Humphreys 2015, p. 25).

In his first book, *Out of Control: The New Biology of Machines, Economic and Social Systems*, Kelly argued that human intelligence is not organized as a central structure but more like a bee-hive of small components (Kelly 1994). Clearly responding to Norbert Wiener’s *Cybernetics; or Control and Communication in the Animal and the Machine* as well as Deleuze’s *Postscript on the Society of Control*, Kelly’s use of biological metaphors to conjure up the formal logic of computational machines can be read as yet another attempt to annul the ontological distinction between the human mind and computers that Dreyfus resisted his whole life (Wiener 1948; Deleuze 1992). Kelly not only claims that computers come close to resembling the functioning

⁶ <http://quantifiedself.com/> (accessed 10 October 2017).

⁷ Stephen Humphreys is Associate Professor of International Law at the London School of Economics (LSE). Arguably, the term ‘datasphere’ was first coined by science-fiction author Dan Simmons in his book *Hyperion* (Doubleday Foundation, 1989). In it, humankind is scattered across the universe after the destruction of planet Earth. In a way that seems prescient of Kelly’s *technium*, Simmons describes how dispersed communities are connected by portals that give access to the “datasphere,” the ubiquitous realm of autonomous artificial intelligences residing within the “technocore” which provides humans with all the information they need for survival and well-being.



Fig. 2 The birth of Sophia, a creation from Hanson Robotics. Credits: still from the documentary 'More Human Than Human' (Tommy Palotta, Femke Wolting), © Submarine Amsterdam 2018

of the human mind, but also that computational machines are becoming more biological, i.e., organic—the explosive growth of ‘machine learning,’ AI algorithms that autonomously learn languages or complex games like chess and Go, seems to prove his point. The *Technium*, or the network of interacting components, creates a self-sufficient technological dynamic that effects humans: biological evolution and technology move forward together (Kelly 2010). In Kelly’s techno-human vision, ever more complex AI technologies will be embedded in everything we manufacture and produce in the near future, while we humans constantly generate data for this new kind of intelligent systems. In this relentless process, as Kelly describes it in his latest book *The Inevitable*, humans integrate more and more digital technology into their lives, and computers and humans will increasingly become co-dependent of one another (Kelly 2016). This, to Kelly, is highly beneficial for the enhancement of nothing less than our own self-realization and thus a tendency to embrace without much hesitation. Andy Clark, another popular scientist in the field of ubiquitous computing and AI, even typifies man as a “natural-born cyborg,” living in a world in which “human-technology symbionts” affect everything, including our sense of ‘selve.’ The term ‘cyborg,’ or cybernetic organism, was coined in the 1960s, but by adding the phrase ‘natural born’ to it Clark stresses the fact that humans have always had a natural inclination to annex, implant, or subordinate to all kinds of technological tools, prostheses, and devices. Similar to Marshall McLuhan’s notion of technologies as extensions of the body and society, (McLuhan 1964), Clark argues that it is no “futuristic mumbo jumbo” to call human beings “natural born cyborgs,” but that it is an existential feature of “our distinctively human nature” (Clark 2013, p. 3) to interact with technology, including today’s computational and data systems (Fig. 2).

Our entry point in this techno-political discourse is the question whether the concept of man as aggregate of data—of which Kelly’s idea of the Quantified Self is an excellent

example—leaves any room for human agency in the good old Greek tradition of knowing and doing yourself. We are certainly struck by the fact that Kelly and Wolf in their popular TED talks and other media appearances do not consider the concept of Quantified Self in a more reflective manner. Their rhetoric is that of tech gurus preaching the blessings of life-changing technologies at a volume that drowns out any critique, thereby potentially using or even abusing their followers’ naivety. In contrast to Clark, they make their views operational.

We are not alone in this critical assessment of Kelly’s thought, however grand and compelling it might be. Technology critic Evgeny Morozov, writer of *Net Delusion* and *To save everything, click here: the folly of technological solutionism*, has already settled his difference of opinion with the QS prophets years ago (Morozov 2011a, 2013). In his review of Kelly’s book *What Technology wants* (2010), Morozov ventilates his fundamental misgiving of Kelly’s “uncritical and laissez-faire approach to technology” (Morozov 2011b). He basically reads Kelly’s broad concept of *Technium*, which the latter describes as the “global, massively interconnected system of technology vibrating around us” (Kelly 2010) as a rehash of German ‘technics,’ that is, an expanded notion of technology that was called for and developed in the early twentieth century to describe the socio-cultural consequences of industrial culture.⁸ He also lashes out at Kelly’s evolutionary rationale and the ensuing dubious morality that inflates the ‘I-era’ to an unprecedented degree—to wit: the majority of QS technologies focus on the individual, narcissistically gathering detailed data about himself. With this, Morozov’s reply to Kelly’s techno-utopian convictions and predictions triggers a debate that is a must read for anyone interested in the pros and cons of data-technological culture.

For what if the new tracking methods are used (or abused) by marketers and planners not so much for assisting the individual ‘Self’ to grow and thrive, but to control collectives of ‘Selves’ as linked atoms in money-making structures called target groups? Also, companies and products, as well as market strategies and sales pitches, can be scrutinized on

⁸ On Kelly’s lack of knowledge of German technics, Morozov writes: “In the early years of the twentieth century, the German debate about *Technik* made its way into America, when Thorstein Veblen discovered some of the key German texts and incorporated them into his own thought. But Veblen chose to translate the German *Technik* as ‘technology,’ most likely because by that time the English word ‘technique,’ the more obvious rendering, had already acquired its modern meaning. To his credit, Veblen’s ‘technology’ preserved most of the critical dimensions of *Technik* as used by German thinkers; and he masterfully located it within contemporary debates about capitalism and technocracy.” Morozov discusses other German thinkers from Simmel to Heidegger who took part in the debate and concludes: “Most of these thinkers posited the growing autonomy of technology—including the self-reinforcing behavior of the system that Kelly emphasizes—and they found this prospect terrifying” (Morozov 2011b).

whether they address us as autonomously thinking individuals or primarily as biddable parts of a market segment. In other words: for the same token, you could see all of the above as the ultimate victory of the economization of our culture, which translates everything that occupies us into transactions based on quantifiable values, and which considers every aspect of our lives as a function of the market. The ‘agora,’ where the ancient Greeks discussed their ideas on authenticity and man as autonomously thinking and acting being, has been narrowed down to a market place in which the main issue is the kind of profit and loss that can be expressed in hard numbers.

It is the victory of the ‘third-person perspective’ of man. Philosopher Jos de Mul reminded us of this concept of his older colleague Helmut Plessner during a debate in the context of the exhibition *The Life Fair* at the New Institute for Design, Architecture and Digital Culture in Rotterdam (2016).⁹ De Mul even suggested to add a ‘fourth-person perspective’—that of being totally immersed in a virtual other, of virtually experiencing being someone else [a kind of new, upgraded version of Sherry Turkle’s early idea of ‘the Second Self,’ by which she first described life on the screen (Turkle 2005)]. The quantified human looks at himself as another and can, with the help of mediating technology, experience himself as such. Though at first sight this seems to neutralize the alienating essence of the human condition, this objectifying perspective, if we take it too absolutely, may carry the risk of distancing us from our authentic self. From this perspective, the actions this self can undertake, its agency, more and more become like handling a machine, and agency is experienced less and less as a sensitive and thoughtful acting based on the kind of complex considerations we call ‘authentic.’ The third-person perspective objectifies our view to ourselves, turns us into a product that can be made, adjusted, improved. It looks at the self and the body as a technological contraption, which can be hacked and onto which other products, services and (social) media can be mounted—a process in which the body and the self that is linked to it tendentially become a mere commodity. During the same debate, artist Simone Niquille wondered what this “colonization of the body by means of technology” would mean for the agency of people. For if “the commodified self” is a self compiled from the offerings of the supermarket of life, then that raises the question of whose standards and values are built into that ‘off-the-shelf’ self: those of the consumer or those of the producer, those of the

individual or those of ‘governmentality’ (Foucault’s term for how governments produce citizens).

Concurrently, this commodification of the self, and of the individual connected with it, triggers a debate on agency and morality that potentially explodes the anthropocentric paradigms that have ruled this discourse for millennia. As David Gunkel has argued in his overview of the recent debate on ethics and machines, *The Machine Question: Critical Perspectives on AI, Robots, and Ethics*, ethics “has been and remains an exclusive undertaking. This exclusivity is fundamental, structural, and systemic.” (Gunkel 2012, p. 160) For centuries, man has been considered an exclusive category—the sole life form capable of considering good and evil, and of acting ethically with regard to the difference. Since the debate on the “*consideranda*” of ethical agency has been broadened to animal rights, environmental ethics and the morality of algorithms, this innate ‘exclusivity’ is under scrutiny. Just like animals can ‘suffer,’ machines can be responsive in ways that prompt moral judgment, which at the very least opens up a debate on what F. Allan Hanson calls a “joint responsibility” of man and machine, where “moral agency is distributed over both human and technological artifacts” (Hanson 2009, p. 94, quoted in; Gunkel 2012, p. 165). Hanson’s “extended agency theory”—in a sense an extension of Latour’s actor–network theory, which we will discuss below—“introduces a kind of ‘cyborg moral subject’ where responsibility resides not in a predefined ethical individual but in a network of relations situated between human individuals and others, including machines.” (p. 165).

5 The technological I

Humans are essentially technological beings, techno-organic hybrids, or “natural born cyborgs,” as technology philosophers, sociologists and information experts such as David Gunkel, Allan Hanson, Bruno Latour, Peter Paul Verbeek, and Philip Agre unwearingly continue proving, beyond Dreyfus’ life-long philosophical reluctance. The colonization that Niquille is referring to has since time immemorial been an interaction between man and machine, between the authentic individual who reflects on his own actions, and the standards and values programmed into his tools. But who is programming man, if the self has become an integral part of the machine, has become a tool itself? If the third-person perspective becomes the dominant perspective onto ourselves; if, therefore, we start seeing ourselves as products—as something mechanical—then this interaction is threatened to grind to a halt. The ongoing discourse on the relation between man and machine then becomes an exchange of data between two essentially technological entities—and the question of moral agency shifts from man towards machine.

⁹ *The Life Fair – New Body Products*, Het Nieuwe Instituut, Rotterdam, 12/06/2016–08/01/2017. This co-authored journal article is a substantially revised and extended adaptation of Max Bruinsma’s essay “Authenticity as Product”, published on the exhibition’s website in November 2016: <https://thelifefair.hetnieuweinstituut.nl/en/authenticity-product>.

The French sociologist and philosopher of science Bruno Latour played a major role in shaping this discourse. Latour, of course, is known for his contribution to the so-called ‘Actor–Network Theory’ (ANT), which posits that everything in this world is interconnected within constantly shifting and interactive networks, in which human agency is just one factor—or ‘actor’—among others, such as physical objects, philosophical ideas, technological inventions, and socio-political processes. All of these actors together determine what Latour calls the ‘social context,’ in an expanded definition of that term (Latour 1991a). Latour’s theory culminates in his book *We have never been modern* (*Nous n’avons jamais été modernes*), in which he offers an alternative view on the modern project, namely that there is no (in fact, there never has been any) clear-cut distinction between the natural and social world, as institutionalized by the separation of the sciences and the humanities (Latour 1991b). Modernity may have wanted to make a clear distinction between the human and the non-human, the natural and social, but reality, the ‘life-world’ is full of hybrids.¹⁰ Latour warns for the dangers of the classic modernist model with his example of the ozone layer: If the natural sciences do not take into account the realities of the socio-political context, with an often reluctant public, stubborn politicians and profit-seeking companies who, in turn, might not want to listen to legitimate scientific warnings because ‘scientists are still debating the problem,’ and everyone just remains within their own bubble, a natural disaster might ensue due to the ever-widening ozone hole—and this is true for many other challenges and concerns in this world, including technological ones (Latour 1991b, pp. 1–10).

Importantly, Latour shifts his philosophical sociology to technological questions in the 1990s. What are the ways in which humans interact with technological objects—the actor–network theory, after all, is a theory on the agency of objects, or more precisely, of “non-humans woven into the social fabric” (Latour 1991a, p. 103). Latour’s turn is already evident in his pivotal essay “Technology is Society Made Durable,” in which he suggests to study sociology in tandem with the history of technology, and poses the question as to whether it is possible to go beyond the divide between sociology and technology. In Latour’s own words: “[We] have to turn away from the exclusive concern with social relations and weave them into a fabric that includes non-human actants that offer the possibility of holding society together as a durable whole (1991a, p. 103).”

Latour’s arguments famously develop around the example of hotel keys, which guests tend to not drop voluntarily

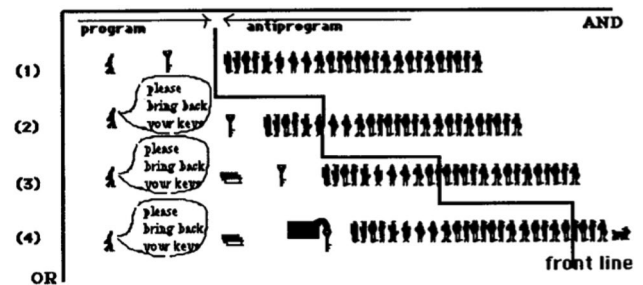


Fig. 3 The hotel manager successively adds keys, oral notices, written notices, and finally metal weights; each time he modifies the attitude of some part of the ‘hotel customers’ group. Source: Bruno Latour, “Technology is Society made Durable”, p. 107

at the front desk before going out. The manager can take different steps to encourage his guests to leave their keys behind, such as asking them politely to do so, putting up a sign or attaching an inconvenient metal weight to the keys so that they want nothing more than to “rid themselves of this bulky object, which makes their pockets bulge and weighs down their handbags.” (Latour 1991a, p. 104) Latour wonderfully captures this in an algorithmic formula, in which he describes the program as the measures undertaken by the hotel manager to encourage the guests to leave the keys, while the anti-program consists of the ways in which the latter undermine that program (not listening, ignoring signs, removing weight, etc.) (Fig. 3).

Although today the ‘hotel key problem’ is virtually annihilated by the introduction of the magnetic key card, which transfers the overview the manager needs of which guests are in and which are out from the material key box to the computer screen, the point of the example is that humans and objects can be seen to interact with each other in a process in which human actors (the hotel manager) are gradually replaced by non-human ones (sign, weight). Since Latour presents this familiar example in the format of an algorithm, it can easily be transposed to digital platforms and other environments driven by data-gathering software, which are inhabited and used by our hybridized, aggregate *Homo Ex Data*.¹¹

However, it is philosopher of computing and AI expert Philip Agre, who most clearly and fully analyses the new forms of algorithmic control built into networked digital systems to track and control people’s behavior. In his remarkable, analytical article “Surveillance and Capture: Two Models of Privacy,” first published in 1994, Agre—a

¹⁰ Husserl’s concept of *Lebenswelt* is taken on by philosopher of technology Don Ihde as ‘life-world,’ in a phenomenology of practically experienced relations between man and technology that inspired Peter Paul Verbeek in his contribution to the *Onlife Manifesto* of 2013 (see note 18).

¹¹ The phrase ‘Homo Ex Data,’ which shows affinity with our idea of man as ‘aggregate of data,’ is the title of an exhibition of 150 Red Dot awarded designs at the Hong Kong Design Institute, 25 November 2017–27 May 2018. See: https://en.red-dot.org/exhibition_homo_ex_data.html (accessed December 10th, 2017).

former professor of information studies at UCLA—already foresees that all spheres of human society might be restructured according to what the author calls ‘the capture model.’ Agre frames this model as an upgraded version of Foucault’s surveillance model, which is relegated to the pre-digital age. Given the importance of Agre’s model for further understanding the society of control (Deleuze), it is worth quoting the information expert here at length:

In naming this model, I have employed a common term of art among computing people, the verb “to capture” (...). The term has two uses. The first and most frequent refers to a computer system’s (figurative) act of acquiring certain data as input, whether from a human operator or from an electronic or electromechanical device. Thus one might refer to a cash register in a fast-food restaurant as “capturing” a patron’s order, the implication being that the information is not simply used on the spot, but is also passed along to a database. The second use of “capture,” which is more common in artificial intelligence research, refers to a representation scheme’s ability to fully, accurately, or “cleanly” express particular semantic notions or distinctions, without reference to the actual taking-in of data (Agre 2003, p. 744).

By comparing and contrasting the models of surveillance and capture in his article, Agre comes to an understanding of the characteristics of the capture model. He clarifies his analysis with a range of examples, from the waiter in a restaurant who captures the guest’s order by a handheld ordering device (first sense of the term ‘capture’—data from inputs), to algorithms that capture human activities that they first structure programmatically (second sense of capture). In all of this, Agre displays an acute critical awareness of the consequences of tracking and capture methods that are built into computer and software design, including concerns regarding unscrupulous corporations using data for the sake of increasing revenues that we expressed above in a critical response to Kelly’s idealistic ideas.

It is Agre’s second sense of capturing activities that is important for an ontological understanding of man as aggregate of data—that the new type of individual is no longer just an undivided, physical body, but is infected by its data representations. Consistent with his phenomenological analysis of the history of ‘disembodiment,’ or the opposition between body and soul, that characterized Western philosophy from Augustine, via Descartes to Turing and Dreyfus, Agre clarifies this representational complex of the capture model through his idea of ‘grammars of action,’ which are programmed into the software: “[T]he phenomenon of capture is deeply ingrained in the practice of computer system design through a metaphor of human activity as a kind of language. Within this practice, a computer system is made

to capture an ongoing activity through the imposition of a grammar of action that has been articulated through a project of empirical and ontological inquiry” (2003, p. 749). User interface design and protocols, in other words, provide users with a range of grammars of action, which actually control and dictate how we act—similar to how grammar dictates the difference between ‘good’ and ‘bad’ sentences. Agre concludes: “As human activities become intertwined with the mechanics of computerized tracking, the notion of human interactions with a computer—understood as a discrete, physically localized entity—begins to loosen its force. In its place, we encounter activity systems that are thoroughly integrated with distributed computational processes” (p. 743). These processes and their ‘grammars of action,’ in turn, are heavily influenced by specific philosophical a priori on the human condition, inscribed into the code and prescribing the range of expected or tolerated action by the user. While Agre’s essay is mostly cited in the context of privacy issues, due to concerns in society about the new type of control that he describes, it is interesting to note that he ultimately shares Dreyfus’ urge for an ontological inquiry into man’s status as human being. As Jethro Masís tells us in his perceptive essay on Agre’s legacy: “Influenced heavily by Dreyfus’ pragmatization of Heidegger, Agre too understands *Sein und Zeit* as providing a phenomenology of ordinary routine activities, and believes Heidegger’s *Analytik des Daseins* can provide useful guidance for the development of computational theories of interaction” (Masís 2014, p. 58).

Latour explores questions of interaction in hybridized relationships between human and technological ‘actors’ in the network society, and Agre methodically thinks these interrelations through on the level of data systems from a critical perspective on “the sedimentation of intellectual history” (Agre 2003, p. 131). It is Peter Paul Verbeek, however, elaborating on the work and tradition of Hubert Dreyfus, Hans Achterhuis, Don Ihde and others, who interrogates the relationship between humans and technology in the context of the themes under scrutiny in this paper: questions of authenticity and individuality in the ontological context of the new concept of man as aggregate of data and the underlying issues of morality that it triggers—what space is left for authentic human agency and individual responsibility in a technologically mediated society?

Verbeek’s first response is that humans and technology develop in tandem: “This you could call the dialectic of technology: we make the technology, but the technology forms us in return. It is a constant leapfrog” (Bruinsma 2015, p. 127). While Verbeek does not deny the potential benefits of Kelly’s “technologies of the Self” he does not share his unmitigated optimism about their effects. With Latour, Verbeek operates from a perspective in which social relations are woven “into a fabric that includes non-human actants,” which implies that the design of any technology that affects

such relations is per se an ethical act that involves moral judgment and critique. As Agre has shown, the grammars of action inscribed into the algorithms that govern our dealings with and within digital technologies are not neutral—they embody a particular history of distancing ‘body and mind’ and a specific phenomenology of how humans experience themselves, suspended between nature and technology. Verbeek advocates a radically transparent attitude towards such biases, which are materialized within technology, in his focus on the way technologies and design mediate them:

[D]esigners should force themselves to be more explicit in their methodology about the implicit, hidden values they project onto future users of their products. These should become readable in the design, just as its function should be. Each product inevitably mediates how people perceive the world, how they behave—and each product inevitably adds bias to this perception. Designers who are ignorant of this fact, or disregard it, are operating in an immoral manner (Bruinsma 2015, p. 130).

Since our “being-in-the world,” as Verbeek quotes Heidegger, is becoming fundamentally affected by new information and communication technologies, we need to develop a new ethics of use regarding these technologies, both in terms of their design and implementation, and in governing the ways in which they are being used socially. Public space is being redefined, and so is the individual’s presence within it. Beyond the rhetoric of unconditioned espousal or hardened antagonism, there is a need for a new ethics that critically assesses the ways in which technologies and humans ‘live together.’ In this context, Verbeek calls for a new “asceticism” in approaching technologies of the self “in such a way that people *want* to be influenced (...) a technology that helps you to consciously engage with yourself as ethical being” (Bruinsma 2015, p. 136), within the social and cultural practices in which you are embedded:

A critical use of information technology then becomes an ‘ascetic practice’, in which human beings explicitly anticipate technological mediations, and develop creative appropriations of technologies in order to give a desirable shape to these mediations. At the same time, the design of information technology becomes an inherently moral activity, in which designers do not only develop technological artifacts, but also the social impacts that come with it. And policy-making activities regarding the implementation of new technologies then become ways of governing our technologically mediated world (Verbeek 2015, p. 224).

Verbeek epitomizes a fine balance between the classical ideal of the authentic human, with his personal freedom of choice, and the ‘society of control,’ with its tendency to

reduce us to our data. Seen from the perspective of Agre’s capture model, Verbeek reminds us of the fact that specific moral codes have always been inscribed in our grammars of action, through law books, religion, cultural norms and mundane phenomena such as door locks. This, again, puts man’s autonomy and authenticity in perspective. Where broadly shared values are concerned, rather than rigorously emphasizing autonomy, Verbeek holds, “[w]e should not shy away from designing stronger ethical influences into products for fear of being regarded as paternalistic” (Bruinsma 2015, p. 137). These values—and their inevitable evolution—on the other hand, should remain a matter of public debate rather than being sealed into algorithms that are impervious to everyone but the coders. For it is in the public arena, that a social morality is developed, which not only guides commonly acceptable behavior in public life, but also deeply affects ideas and feelings of good and bad consciousness in the private realm—this is a key reason for the growing public concerns over privacy issues. Summarizing the effects of data-driven networks on the individual’s stance within the public realm, legal expert Stephen Humphreys remarks that the private–public distinction, which has governed (legal) relations between the powers that be and the individual since Hobbes’ times, “appears to be dissolving” (Humphreys 2015, p. 25). In the wake of this process, the notion of conscience, of reflecting on matters of good and evil and acting according to one’s own autonomous judgment, takes on a new urgency in a globalized world in which the “public sphere” is rapidly turning into a “datasphere”¹²:

And so, to come full circle, it seems conscience is reviving in the datasphere. Indeed, we are poised between the return of bad conscience and of good conscience. On one hand, the paralyzing realization that our thoughts are naked and legible before an omniscient authority, whose knowledge and motives remain inscrutable. On the other, a slow awakening into a context and capacity to act autonomously on the information available to us. (Humphreys 2015, p. 28).

6 Conclusion

The disconcerting aspect of a concept of man as aggregate of data is that it implies that man can only function within the categories—or grammars of action—which enable the

¹² In his paper, Humphreys refers to “... this condition of data immersion as life in the ‘datasphere’. I use this term by analogy with the old ideal of a ‘public sphere’, in order to capture the degree to which data saturation is today a profound, structural aspect of our working, playing, communicating, politicking, networking, consuming, self-policing selves” (Humphreys 2015, p. 5).

processing of these data. This again could turn us into pawns on the chess board of the institutions who formulate these categories, as technical axioms rather than open values that can be discussed publicly and collectively. Verbeek, too, stresses the necessity of openness in the design of technologies that deeply affect the public realm, along with a more thorough understanding of the ways in which these technologies not only facilitate, but basically constitute our dealings with the world: “If human practices and experiences are always technologically mediated, there does not seem to be an ‘outside’ position anymore with respect to technology. And if there is no outside anymore, from where could we criticize technology?” (Verbeek 2015, p. 222).

In our context, one can reverse the above position, and suggest that if everything has become ‘outside’ (as a result of the third-person perspective on man and his self-awareness as aggregate of data incarnate), the authentic ‘inside’ position of an individual who critically reflects on his stance vis-à-vis the ‘outside’ is fatally weakened. Thus, knowing thyself, today, calls for an intensified reflection on the individual’s responsibility with regard to his technological alter ego. An active pondering of the dizzying condition of the individual as quantum rather than atom, constantly switching between his authentic ‘first-person perspective,’ his humanly insufficient ‘second-person perspective’ (his conditionally deficient ability to empathize with others), and the ‘third-person perspective’ that allows him to externalize himself as commodity. If we don’t reflect on what this means to us, humans, and the world that we build together—technology and all—we threaten to lose sight, not only of the distinction between man and machine that was so dear to Dreyfus, but of the space for authentic reflection and agency that until now has been a fundamental characteristic of *Homo Sapiens*. If we evolve into *Homo ex Data*, our “normative framework” (Verbeek 2015 p. 226) needs to evolve with it, in acknowledgment of our condition as human–technological hybrids.

That, in this setting, a deeper awareness, knowledge and critique of AI is urgent, both in academia and in the public realm, is evident from the fact that AI is becoming increasingly ubiquitous, with an increasing impact on our daily lives. One has only to read the news these days, in newspapers, journals, TV shows, documentaries and films, online and offline, to witness the growing public recognition of this development—and a growing unease concerning its consequences. To mention only a few topics which were widely discussed in the mass media during the last months of 2017: Robots are predicted to replace humans in the workplace and radically change the economy and social life—recently a robot has been granted citizenship of Saudi Arabia; the driverless car promises to fundamentally change traffic, triggering concerns on safety and regulation; and there are ongoing and growing worries about security and privacy issues in a world in which humans increasingly become dependent on

connected networks, services and smart devices. The turmoil that arose in the first months of 2018 around the British–American data analyzing firm Cambridge Analytica and its illegal use of millions of Facebook profiles for manipulating elections is another case in point.¹³ This globally discussed affair triggered a substantial loss of confidence in Facebook’s ability and willingness to protect the data of its bank of billions of ‘dividuals’ (and a plummeting of its stock market value), and caused the bankruptcy of the perpetrators, Cambridge Analytica.¹⁴ But equally important from our perspective, the scandal invigorated the global debate on the manipulative potential of data-mining technologies, and the almost complete lack of transparency on how their “capture models” and “grammars of action” not only rule analytic procedures, but represent actual agency IRL (“in real life”).

All of this reinforces the idea that has settled in professional and academic circles already for some time, that AI will have a deep impact, not only on technology and science, but on the ‘life-world’ at large. It underscores the urgency of studying the ways in which people interact with AI systems, from a human-centered approach. If our grammars of action are written solely from the viewpoint of the computer’s protocols, we run the risk of becoming ensnared in ever-tightening algorithmic straightjackets that control rather than facilitate our lives, and foreclose the optimal conditions for a well-functioning, democratically organized “telematic society” that Vilem Flusser already advocated in the 1980s, with the lived-through historical traumas of the grand totalitarian regimes in the twentieth century still in his mind (Van der Meulen 2010, pp. 206–207).

The fundamental redesign of public space and the private realm is too important to be left to the coders and the institutional and commercial interests they serve. When I, an individual human being, am addressed as aggregate of data, my authenticity is put into question. If Google, Facebook, Twitter, YouTube, (and every other digital platform that I happen to feed with traces of my existence) own my data—which they do, because in clicking “agree” to their conditions of use, I surrendered them to them—how do I own myself? To answer this rather existential question, we may refer back to the old Greeks, to Kant and to Dreyfus by formulating this maxim, which, ultimately, also infers what computers should not do: I cannot and will not delegate my own authentic consideration of good and evil to the technologies, institutions

¹³ See a comprehensive overview of the media coverage by the british newspaper The Guardian: <https://www.theguardian.com/news/series/cambridge-analytica-files/all> (accessed May 11th, 2018).

¹⁴ On their website, the data-mining firm is quite candid about their aims: “Data drives all we do. Cambridge Analytica uses data to change audience behavior.” See <https://cambridgeanalytica.org/> (accessed May 11th, 2018).

and interests of which I am an integral part, but into which I do not want to dissolve completely.

Acknowledgements We would like to thank our friend Prof. dr. ir. Remko Scha (1945–2015), former professor of computational linguistics at the University of Amsterdam, for his pioneering, always inspiring, and exceptionally creative teachings, thinking and artistic practices in computational science and algorithmic art. Remko rightly advised to take the work of Hubert Dreyfus seriously.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Adorno TW, Horkheimer M (2002) The culture industry: enlightenment as mass deception. In: Schmid Noerr G (ed) *Dialectic of enlightenment* (trans: Jephcott E). Stanford University Press, Stanford, pp 94–137
- Agre P (2003) Surveillance and capture: two models of privacy. In: Wardrip-Fruin N, Montfort N (eds) *New media reader*. MIT Press, Cambridge
- Augustine (2010) *On the free choice of the will, on grace and free choice, and other writings*. Cambridge University Press, New York
- Bruinsma M (2015) Designers do not determine what a thing does. Interview with Peter Paul Verbeek. In: *Design for the Good Society*, NAI010, Rotterdam, pp. 125–137
- Clark A (2013) *Natural Born Cyborgs: minds, technologies, and the future of human intelligence*. Oxford University Press, New York
- De Souza e Silva A (2006) Interfaces of hybrid spaces. In: Kavoori AP, Arceneaux N (eds) *The cell phone reader*. Peter Lang Publishing, New York
- Deleuze G (1992) Postscript on the societies of control. *October* 59:3–7
- Dreyfus H (1965) *Alchemy and artificial intelligence*. The Rand Corporation, Santa Monica
- Dreyfus H (1993) What computers *still* can't do. A critique of artificial reason. MIT Press, Cambridge (**second printing**)
- Dreyfus H (2001) *On the Internet*. Routledge, New York
- Gunkel DJ (2012) *The machine question: critical perspectives on AI, Robots, and Ethics*. MIT Press, Cambridge
- Hanson FA (2009) Beyond the skin bag: On the moral responsibility of extended agencies. *Ethics Inf Technol* 11:91–99
- Humphreys S (2015) Conscience in the datasphere. *Humanity* 6(3):361–386 (**LSE Legal Studies Working Paper No. 11/2015**). Available at SSRN: <https://ssrn.com/abstract=2613310>. Accessed 29 June 2018
- Kant I (1998) *Groundwork of the Metaphysics of Morals*. (trans: Gregor M). Cambridge University Press, New York
- Kelly K (1994) *Out of control: the new biology of machines, economic and social systems*. Basic Books, New York
- Kelly K (2010) *What technology wants*. Penguin Books, London
- Kelly K (2016) *The Inevitable: understanding the 12 ecological forces that will shape our forces*. Viking, New York
- Latour B (1991a) Technology is society made durable. In: Law J (ed) *A sociology of monsters essays on power, technology and domination*. Sociological Review Monograph, vol 38, pp 103–132
- Latour B (1991b) *Nous n'avons jamais été modernes. Essai d'anthropologie symétrique*. Le Découverte, Paris
- Luther M (1957) *The bondage of the will: a new translation of de servo arbitrio* (1525). (trans: Packer JL and Johnston OR). Fleming H. Revell Co., New Jersey
- Manovich L (2008) *Software takes command*. http://softwarestudies.com/softbook/manovich_softbook_11_20_2008.pdf. Accessed 29 June 2018
- Másís J (2014) Making AI philosophical again: on Philip E. Agre's legacy. *Continent* 4(1):58–70
- McLuhan M (1964) *Understanding media: The extensions of man*. McGraw Hill, New York
- Morozov E (2011a) *Net delusion: the dark side of internet freedom*. Public Affairs, New York
- Morozov E (2011b) *E-salvation: Kevin Kelly's What technology wants*. *New Republic* (digital edition), March 3. <https://newrepublic.com/article/84525/morozov-kelly-technology-book-wired>. Accessed 29 June 2018
- Morozov E (2013) To safe everything, click here. *Technology, solutionism, and the urge to fix problems that don't exist*. Penguin Books, London
- Nietzsche F (1886) *Jenseits von Gut und Böse*. CG Naumann, Leipzig
- Nietzsche F (1967) *On the Genealogy of morals, Ecce Homo*. (trans: Kaufman W). Random House, New York
- Sartre JP (1945) *Huis Clos, suivi de Les mouches*. Éditions Gallimard, Paris
- Turkle S (2005) *The second self: computers and the human spirit*. MIT Press, Cambridge (**First published in 1984**)
- Van der Meulen S (2010) Between Benjamin and McLuhan: Vilem Flusser's media theory. *New German Critique* 110:181–207
- Verbeek PP (2015) Designing the public sphere: information technologies and the politics of mediation. In: Floridi L (ed) *The online manifesto—being human in a hyperconnected era*. Springer, Cham
- Wiener N (1948) *Cybernetics; or, control and communication in the animal and the machine*. MIT Press, Cambridge