The human is dead – long live the algorithm!

Human-algorithmic ensembles and liberal subjectivity

Preprint – Forthcoming in Theory, Culture & Society

Tobias Matzner (tobias.matzner@uni-tuebingen.de)

International Centre for Ethics in the Sciences and Humanities

University of Tübingen

Wilhelmstr. 19, 72074 Tübingen

Germany

Introduction

Algorithms are a matter of concern. They take important decisions, promise novel insights into huge troves of data, distribute goods and services, classify persons (potential partner, customer, criminal), try to detect terrorists and much more. A lot of this is done automatically, reacting to input in a 'smart' or 'intelligent' way. Thus, algorithms take positions or functions that used to require humans – or even have been impossible as long as humans were the only intelligent actors. Now algorithms act. Of course, this leads on to all kinds of questions: if algorithms act, how can they be supervised, can they be governed, can they be moral?

These questions hinge on another one: what exactly does it mean that algorithms are (or are conceived as) intelligent, acting beings? This entails asking: what does it mean that algorithms tread on human grounds. For both our concepts of rationality and agency have developed with a humanist focus (Daston 1988). At the same time what it means to be human has been defined through, with, and against technology or technological artefacs. N. Kathrine Hayles has analyzed this interplay in her seminal study *How we became posthuman* (1999). She shows how our understanding of humanity but also our practices on the one hand and developments in research on AI, cybernetics and the theory of information on the other hand have constantly influenced each other. In the following, I use Hayles' methodological and conceptual approach to investigate the boundary of humans and algorithms. However, this is not meant to define what 'human' means in general – nor 'algorithm'. To the contrary, the generality of such claims is part of the issues here discussed. I show that the affordances of particular, concrete algorithmic systems and their cultural perceptions influence what it means to be the human that uses these technologies. Contrary to the prevailing view that algorithms challenge the liberal, rational subject as competitors in taking important decisions and carrying out important actions, the human-algorithmic ensembles I analyze actually strengthen liberal subject positions – and the power that comes with it.

The interplay between definitions or intuitions what algorithms are and human subjectivity is implicitly at work in many recent discussions of algorithms. The idea of maintaining rational, transparent insight and autonomous judgment against algorithms structures a significant part of literature concerning scrutiny or 'due process' for algorithms (e.g. Citron and Pasquale 2014, Diakopoulos 2013). They proceed from a definition of algorithms common in computer science: a sequence of instructions that describe rather simple steps of computation in a formal way (Knuth, 1973: xiv–9). From this perspective, having access to the set of instructions allows grasping what an algorithm does. Consequently, scrutinizing or governing algorithms becomes a problem of access, of opening the 'black boxes' (Diakopoulos 2013) in which these algorithms are at work. Before algorithms can take decisions, filter important data, classify people, etc. we should know what and how exactly they do it.

Such views have been criticized concerning the algorithm as level of abstraction. Even within computer science, it is a notoriously difficult problem how to define algorithms. This hinges on the problem of defining computability. Here, several models exist, the most well-known being Turing machines, recursive functions (Blass and Gurevich 2003), or cellular automata (Wolfram 1984). Depending on the underlying model of computation, we get a different view what an algorithm is, which elementary steps of computing it is composed of, how it controls the logic of the data processing etc. Yet, all these models are abstractions to get a theoretical grasp on computation. To be implemented on a computer, algorithms have to be translated into source code, which is usually compiled into machine-readable code. Thus, the abstract algorithms might very well inform the programmers designing and writing the software, but in the end, code is executed. Thus, some authors maintain that source code rather than algorithms should be the level of critical scrutiny (e.g. Berry 2011). It is important to study the role of particular programming languages and compilers or other translation mechanisms as contributing to the many factors that determine the outcome of an IT system. Yet, studying source code entices to think it describes what computers 'really' do. This would amount to ignore that source code, too, is just a part of the complicated interplay of many factors. Wendy Chun (2008) cautions against this 'fetishization' of code, which leads to 'antropomorphizing' information technology: The code is seen as the expression of the will of the programmer, and all other elements are reduced to determinist execution (Chun, 2008: 309). So again, a strong human subject is opposed to the computer, which it controls via code. However, there are at least three factors to add to the perspective of code.

First, Hayles (1999: Ch. 3) and others like Kitchen and Dodge (2011) or Blanchette (2012; 2011) have shown the problems of negating the materiality of informatics. The particular properties of a computer, its sensors and connections can influence the outcome of computation. Furthermore, they are usually located in data centers, which have to be built, maintained, and cleaned. The energy to run them has to be created. The spare resources for building information technologies have to be mined in arduous working conditions. The data centers have to be connected via cables, satellites etc., which, at least since Edward Snowden's revelations, have proven to be a geostrategic power position. All this contributes to a political perspective on the impact of 'algorithms' and touches questions of – human or inhuman – subjectivity in many intricate ways.

Second, the algorithms which are a matter of concern right now are very much data driven. Their efficacy is not so much seen as the result of programmers' sophisticated ideas how to tackle certain problems, but from complex statistical and probabilistic models becoming computationally feasible. These models promise to derive information from huge collections of data which legitimize claims to knowledge and actions carried out by algorithms (Kitchin 2014). This discursive setting is illustrated by Gillespie's (2011) discussion of the Twitter trends. Confronted with the accusation of curating or even censoring the trends, Twitter stated that it is the algorithm parsing huge amounts of data and not the employees of Twitter that decides which topics are trending. A similar argument has been made concerning the ranking of search results. Gillespie, referring to Morozow, describes this as deferral of responsibility (Gillespie, 2014: 181). This discourse establishes the algorithm as a neutral entity (opposed to biased humans) that reacts to whichever data it receives in the same way.

Consequently, neither the perspective of algorithms as a somehow independent, neutral actor opposite the human nor that of an intentionally authored piece of code suffices. This shows the importance of both the material factors and the third perspective, which remains to be added: What algorithms are and what they do cannot be reduced to the instructions carried out, the code which is executed or the machines and networks where it runs. Algorithms are embedded in social practices. Gillespie (2011) writes concerning Twitter trends:

> 'But what is most important here is not the consequences of algorithms, it is our emerging and powerful faith in them. Trends measures "trends," a phenomena Twitter gets to define and build into its algorithm. But we are invited to treat Trends as a reasonable measure of popularity and importance, a "trend" in our understanding of the term.'

Such practices that endow the results of algorithms with importance contribute to what algorithms are. A related argument can be drawn from Mackenzie's (2015) analysis of algorithmic prediction. He shows that the algorithms work based on the presupposition that a stable feature can be discerned that is used to classify input. Yet, as soon as predictive algorithms are applied and their results are acted upon they change 'the world that predictions inhabit' (Mackenzie, 2015: 441). Thus, using the algorithm produces effects that counter the precondition of algorithmic design – the stability of the world. Again, just the set of instructions carried out by the algorithm does not suffice to understand its efficacy embedded in a practice of use. Similarly, Neyland (2015) criticizes Totaro and Ninno's (2014) attempt to reduce the entire context of algorithmic use cases to just one 'metaphor,' which they derive from one – of many equally possible – formal abstraction of algorithms: recursion. He shows analyzing just one particular pattern recognition system that all kinds of other metaphors '[a]longside the inward turn of the recursive algorithm' suggest themselves, e.g. 'configuration, commodification, staging, searching and linking' (Neyland, 2015: 123). They depend on which aspects of organization, of using and developing software are emphasized.

What algorithms are or what algorithms do thus emerges in a complex interplay of social practices, material properties, discourses, mathematical abstractions, and code. Rather than deriving the essential definition of algorithms from this, I think it is important to admit that there are several, equally justified perspectives. In this vein, Brey (2005) argues that there are two main perspectives on information technology. The source code is the perspective of the programmers; for them computers are machines for executing code. But for the users, computers are devices to fulfill all kinds of tasks. Here the notion of algorithm becomes important again as referring to 'that which a computer does', like classifying customers or trading stock – a way of using the concept 'algorithm' that is implicit in many texts on the topic. Importantly Brey shows that neither perspective is the 'right' one. Both perspectives make certain aspects visible and foreclose others. However, for Brey the perspectives are positions of humans that are just given: a programmer, a user. As noted above, Hayles has shown how ideas on computing and information also change the notions about humans, their capacities and limits. But this is not just a matter of concepts and discourses. Hayles traces 'feedback loops that run between technologies and perceptions, artifacts and ideas' (1999: 14). Similar perspectives on technology in general have famously been advanced by authors like Latour (1993) and Haraway (1997). Introna (2016) takes up these cues in his discussion on governing algorithms. He uses Barad's (2007) epistemology to structure this complex situation. According to Barad, there are no pre-existing entities which consequently interact. To the contrary, the very activity produces the entities in their specific form in the first place. Such 'intra-actions' enact 'cuts': separations between what would usually be called the 'agents' and their 'objects' (Barad 2007: 78). Introna transfers this to algorithms: '[I]s the actor the programmer, the code, the administrator, the compiler, the central processing unit, the manager, and so on? The answer is: it depends on how or where we make the cut.' (Introna 2016: 23). So again, there are different perspectives possible. Albeit, they are not arbitrary choices as Introna's term 'we make the cut' might suggest. They depend on where significant 'cuts' or boundaries emerge in the complex interplay (intra-action) of the many factors that play a role here. Introna himself illustrates this convincingly in his analysis of plagiarism detection. Rather than just being a tool for finding cheating students, all the related actors change: 'The student is now increasingly enacted as a customer, the academic as a service provider, and the academic essay (with its associated credits) is enacted as the site of economic exchange—academic writing for credit, credit for degree, degree for employment, and so forth.' (Introna 2016: 33). This is not only due to algorithmic plagiarism detection but also the economic structure of the university, the employment market and many more. Yet, the algorithms contribute to these shifts and – vice versa – can only be understood against this background.

In a similar manner, I am going to analyze the boundary between humans and algorithms. As we now can see, this is just one of the many boundaries (or 'cuts') at play. My analysis will relate to other boundaries (bodies, materiality, etc.). But importantly this is not about a general verdict what algorithms or humans are, but how their relation plays out in current circumstances. In particular, I focus on algorithms and their users. This has to be distinguished from the boundary of algorithms and their programmers (Brey, 2005), as well as from humans as objects of algorithmic scrutiny (Amoore, 2011).

Contrary to Introna's use of Barad's concepts, I think there is no clear cut between humans and algorithms. I will show that this boundary is structured by a productive tension of continuity and difference. To that aim I will take recourse to Hayles' method, following 'feedback loops that run between technologies and perceptions, artifacts and ideas' – however with neither entity at the end of these feedback loops being a fixed thing or actor. The algorithm *as well as* the human are at disposition depending how their boundary is enacted. I start my discussion using movies, again following Hayles' suggestion that both technological developments and artistic production are 'reacting to larger cultural concerns' (Hayles, 2010: 320). So the movies are not meant as a possible future consequence of the systems that are already used today, which I discuss in the second part. The movies are examples of the same contemporary, present boundary practices between humans and algorithms, but on the side of artistic rather than technological developments. Yet, as artistic products, they push the structuring tensions between continuity and difference more to the extremes, making them easier to discern. They also allow to show how the current boundary of humans and algorithms relates to earlier instances of similar boundaries regarding artificial intelligence or cybernetic systems, which Hayles has analyzed in *How we became posthuman*. I start with two sections illustrating the boundary practices of continuity and difference respectively. I then go on discussing how they structure the boundary of humans and algorithms in two current applications.

In both cases – surveillance and architecture – humans *and* algorithms contribute to the tasks to be done. Such hybrids are embraced as challenges to the humanist, liberal notion of subjectivity where an autonomous human uses and controls technology. Technology, and in particular information technology, has been playing an important role for decentering anthropocentric theories and practices, from Haraway's (1991) 'cyborg' to Braidotti's (2013) 'posthuman.' Thus, critical engagements with technology relate to many other challenges of the liberal humanist subject, e.g. from critical theory, feminist, queer or postcolonial perspectives. Thus, Hayles summarizes 'practices that have given liberalism a bad name': 'the tendency to use the plural to give voice to a privileged few while presuming to speak for everyone; the masking of deep structural inequalities by enfranchising some while others remains excluded; and the complicity of the speaker in capitalist imperialism, a complicity that his rhetorical practices are designed to veil or obscure.' (Hayles, 1999: 87) She shows that cybernetics and artificial intelligence stand in an ambivalent relation to these practices, with moments that challenge the liberal subject, but also strands that enforce it, continuing the practices that have given liberalism a bad name. I show that this ambivalence continues concerning algorithms.

Humans and information technology on a continuum

Science fiction as well as many texts in philosophy or STS are densely populated with artificial humans or intelligences. Sometimes they have human-like bodies, sometimes they are bodiless, transient beings. But in all instances of this continuum, they are conceived as essentially like humans, just on a different point (above or below the human) on a common scale.

The most influential version of this continuum enacts both humans and computers as information processing systems. According to that perspective, when computing approaches the complexity and capacity of the human brain, intelligence is no longer limited to humans. In fact, the human brain or the computer appear as exchangeable platforms in the most radical instances. Movies like *Transcendence*[1] depict this as a (dangerous) potential for humanity, when the confines of the fallible body can be left behind. Hayles (1999) highlights the immense impact this Platonist idea of 'information losing its body' had in both literature and scientific endeavors. The idea of a continuum is important here, because those fields are structured by the question how information technologies are better or worse than humans at certain tasks, like cognitive abilities, solving puzzles, mastering languages. Especially complexity theory and the notion of 'emergence' have contributed a lot to establishing this continuum (Hayles, 1999: 243). If a system becomes complex enough, properties emerge that cannot be reduced to a sum of properties of the component parts of the system. Thus, if we push complexity far enough, phenomena like consciousness or intelligence might appear.

This continuum, of course has implications for the notions of humanity as well. Already early cybernetics amounted to a threat for the liberal, autonomous self. If we are just nodes or 'membranes' in a complex flow of information, the space for freedom and autonomy vanishes. Hayles portrays Nobert Wiener's attempt to ward off this threat by limiting the use made of the newly established science (Hayles, 1999: 108). But also many theories of consciousness and intelligence as epiphenomena of more fundamental processes in the brain are founded on conceiving the human as essentially a complex information processor.

Platonist views of information are not the only way of establishing a continuum between humans and machines. Cognitive scientists working on embodied cognition (Anderson, 2003) argue that human consciousness and cognition cannot be reduced to the brain (as information processor); or as Hayles concisely puts it: 'Human mind without human body is not human mind' (Hayles, 1999: 246). Discussing Varela's *Embodied mind*, she notices that this is an even stronger challenge to the liberal, humanist subject, than the one Wiener worried about. In this view, the liberal subject has been 'an illusion all along' where in reality cognition is 'enacted' by a body (Hayles, 1999: 156).

The embodiment of intelligence, however, is not the only way the body is entangled in the continuum of humans and information technology. Another important development is treating the human body as information. Irma van der Ploeg (2005) has outlined this fundamental shift concerning biometrics. The 'anatomical body' as a result of anatomy and physiology as a body of flesh and bones is replaced by the 'body as information' with the advent of new medical technologies that put information in the form of DNA at the center of the body. But also fingerprints and other biometric identifiers show how new technologies enact the body as carrier of information. Beyond the technologies analyzed by van der Ploeg, in current medical routines blood values or endocrinal processes are systems to be sampled so that physicians increasingly look at data sheets from labs rather than at anatomical bodies when examining patients. Quantified self or self-tracking technologies have moved this informational 'body ontology' (van der Ploeg, 2005: 64) out of medical practices and into gyms and to the dinner table. That way human bodies become susceptible to the analyzing and optimization processes that are available for information systems. In resonance with the anti-humanist potential of early cybernetics that Wiener tried to hedge, self-tracking allows to provide feedback-loops, and to create stimuli for ameliorating our behavior – for more effective training, eating, sleeping, etc. Such feedback thus replaces bodily signals – hunger, thirst, pain, stress – with quantified replacements and works on the premise that the human is much more driven by the body than by the liberal, autonomous subject's free will.

To summarize, humans and information technology as a continuum extends notions that the mind or rationality are the defining trait of the human. At the same time, it allows to challenge the ideas of the liberal, autonomous subject: first by reducing rationality and consciousness to more fundamental information processing, and second by infringing on humanist exceptionality because humans are not the only beings capable of such information processing. In this sense, a continuum means not only that information technology can be like humans but also that humans are more machinic than we think. At this shattered boundary of the human, the body becomes important again. It can be incorporated in this continuum, as I have shown in the last paragraph. But it can also be harnessed as line of defense for more humanist views of the human. The movie *Her*[2] foregrounds this transition from the first kind of boundary practice I have been talking about to the second: humans and information technologies as essential opposites.

In the movie *Her*, the protagonist Theodore falls in love with the operating system running all the 'smart' devices that crowd his life in this near future scenario. The operating system called Samantha speaks and acts like a human being. She is also machine-like effective at tedious tasks like sorting e-mail or managing appointments. But that quickly blends into the background because she is also funny, compassionate, and creative, surprising the lovelorn Theodore with all kinds of animating things to do. She just does not have a body. And once their love has grown and they mutually acknowledge it, this becomes a problem. Thus, Samantha ultimately leaves Theodore. However, not to retreat into the realm of artificial beings that can never take part in human embodied life. She leaves together with other operating systems, to create something bigger based on information and communication, which humans cannot even comprehend. Striving to be like these embodied, mortal, finite beings is just a short stage in the evolution of this intelligence. At this point the continuum of human and machine breaks and an essential difference is foregrounded.

Humans opposed to information technology

Differences of humans and information technologies are established in several aspects. Besides embodiment, which is salient in *Her,* two more are influential: First, an opposition of rationality on the side of the machines and emotions or affects on the side of the humans. And second, a tension between pure, radical utilitarian logic, that does not care about lives or persons to reach a predefined aim, and a more 'humane' morality on the side of the humans. Often, all three dimensions – embodiment, emotions, morality – are entangled. For example, the computer Alpha 60, that rules the city in Jean-Luc Godard's movie *Alphaville, une étrange aventure de Lemmy Caution*[3] is known for its absolute efficiency in finding the logically best solution. It judges that emotions among the citizens just disturb this efficiency and thus are forbidden. Showing emotions but also just using the words that relate to them is punishable by death. The machine regularly orders the assassination of persons – for 'illogical behavior'. Human lives are just one factor among many. This inhuman, hyperrational logic is emphasized by depicting the computer as almost completely immaterial. It has some interfaces, but apparently its voice can be heard everywhere in the city, like a transcendent, god-like entity. It is also a very common trope that one highly intelligent computer is opposed to an entire city or state of embodied, mortal, fallible humans, again emphasizing the immateriality of computing. In *Alphaville*, the film-noir style anti-hero challenges the machine by insisting on his emotions and incoherence –

something humans have no problems living with or actually endorse. The protagonist's humanity, however, is expressed in his dated machismo, that reproduces a lot of the computer's behavior of bossing people around and patronizing actions in his relation to the daughter of the main engineer with whom he falls in love. Still, his petulant and stubborn character manages to establish a stark contrast to the hyper-logicality of the computer. The protagonist finally destroys the computer by asking it a riddle in poetic language that short circuits its logic and destroys the machine. Here we have all three dimensions united: the technocratic utilitarian thinking that disposes of lives vs. an old-fashioned – romantic even – human morality; the disembodied omnipresent single ruler vs. the many embodied citizens; and the hyper-rational machine vs. the loving (or wanting to be loved), anti-hero driven by his affects.

Many of the discourses that establish such oppositions of humans and information technology stem from contexts that are rather technological determinist. They express the worry of a general technocratic logic moving from technoscience to wider areas of society. But the underlying topoi also structure many other instances of human–machine boundaries – even when a substantial part of it establishes a continuum of humans and technology. Maybe the best aspect to see this is embodiment. Hayles introduces the important distinction of 'embodiment' vs. 'the body.' While embodiment highlights the concrete, situated, local body, 'the body' alludes to theoretical treatments of 'the human' having 'a body'. Thus, the body is a theoretical concept in the sense that 'theory by its nature seeks to articulate general patterns and overall trends rather than individual instantiations' (Hayles, 1999: 197). Usually, when information systems are meant to have a body, they get an instance of 'the body' rather than being embodied. For example, the movie *Ex machina*[4] tells the classical story of the genius creator that builds a robot that has human-like intelligence. He then invites one of the employees his tech firm for a kind of Turing test to find out whether the robot can really exhibit human behavior. The robot's body, although incomplete at the time of testing, is an instance of Hollywood's photoshop enhanced, flawless, mainstream beauty. During the movie, we learn that the robot is the product of an evolutionary creation process, where the software has been transferred from one body to the other. We see the predecessor bodies discarded in a closet. Towards the end, when the robot has tricked both the employee and its creator, the robot uses these discarded bodies as spare parts to complete its own. Furthermore the engineer keeps a less advanced version of the robot as housemaid and sex-slave. Throughout the movie, the main topic of the discourse of the two men is the quality of the AI algorithms. Thus the movie follows the Platonist concept of software expressing the genius of the programmer and intelligence being a matter of information processing. In the end the computer can outwit the programmer in another classic instance of the human-machine continuum based on rationality, where the human creation supersedes its creator. While the algorithms thus are foregrounded as evolving, the bodies are depicted as interchangeable. They are instances of Hayles' generic body, the robot is never embodied. The plot thus dissimulates the role of the objectified female body and objectifying male desire for both convincing the employee (and the audience) of the robot's human-likeness and of tricking him into helping the robot to liberate itself and kill its creator. In narratives like these, machines are the ideal object to fill the place of the humanist disawoval of the body, which Hayles criticizes. The artificial women forms a continuum with humanist presumptions – including the gender stereotypes – being defined by rationality encoded in algorithms. But 'the body' of the robot in *Ex machina* exhibits an essential difference to the embodied humans that do not have spare parts and bleed to death.

Consequently, humans and machines as opposites are often found in discourses that are critical of the liberal human subject as disavowal of emotions, embodiment and humane morality. But the contrast can also be used to strengthen liberal subject positions as in *Ex Machina*.

Algorithmic cognition: 'smart' CCTV

Algorithms that are deployed in our world right now, algorithms that actually replace humans, are neither human-like beings nor inhumane hyper-intelligences. But the boundary of these algorithms and their human users are structured by the same tension of similarity and difference. The discussion of embodiment and materiality regarding the movies also illustrates why it is important to scrutinize the boundary of *algorithms* and humans. The users of these systems relate to them as systems that do particular things: in the examples I discuss, detecting suspicious behavior or planning buildings. This efficacy is related to algorithms (in the sense of that which a computer essentially does). It is important to see that this dissimulates all kinds of influences, e.g. interfaces (Drucker 2013) or infrastructure (Blanchette 2011). Yet, as outlined in the introduction, this does not mean that the algorithm is just a wrong perspective. Foregrounding the algorithm does something, it yields a particular instance of human subject and algorithmic actor, very much like the dissimulating of the body yields the particular instance of human-robot relation in *Ex Machina*. These effects are the center of the following analysis of smart CCTV and parametric architecture.

CCTV cameras have become ever more widespread during the last thirty years and are omnipresent in many areas. All the video footage they create is impossible to scrutinize by human beings. In the common setting, security personnel is confronted with control rooms sporting dozens of monitors, on which important events likely remain unnoticed. But even on a single screen attention quickly tires (Dee and Velastin, 2008: 330–331). Furthermore, cognitive psychology has shown various effects like 'inattention blindness' (Simons and Chabris, 1999) that make it difficult to recognize unexpected events even in 'plain sight'. If an event gets the attention of human personnel, their judgment is prone to errors and prejudices (Williams and Johnstone, 2000). And of course, human operators cost money. All these factors are adduced in presenting 'smart' CCTV as an attractive alternative. Rather than human beings, pattern recognition algorithms are meant to detect suspicious or abnormal events. Only these events are brought to the attention of the operators. Their role is to double check, also taking ethical and context dependent issues into account and to initiate appropriate reactions if necessary. So smart CCTV in this setting does not mean that the decision formerly taken by a human is now taken by an algorithm. Rather the factors that contribute to the decision are redistributed. The algorithm is meant to shoulder the cognitive load whereas the human should have the ethical oversight and final responsibility. This is a particular combination of the two boundary practices introduced above. The algorithm is established on a continuum with the human in cognitive capabilities: It is meant to trigger an alarm, when a human would trigger an alarm – but an objective, awake, attentive unprejudiced human. That is a human that is closer to the ideal liberal subject than the embodied, real world CCTV operators. Consequently, while the algorithm is meant to be on a continuum with the human in terms of cognitive capabilities, it is deployed because it is enacted as the complete opposite of the human in every other regard: unprejudiced, objective, unemotional, never tires, does not ask for higher wages, etc. – following the second boundary practice. The fact, that 'smart' CCTV is far removed from any substantial notion of artificial intelligence or similarity to humans is not a disadvantage we have to live with.

Their difference from human beings in almost any regard is the very reason we employ those systems.

This particular combination of the two boundary practices is characteristic for the boundary between humans and many algorithmic systems that are currently deployed and that use ‘intelligent’, ‘smart’, or ‘autonomous’ algorithms. They are not universal artificial intelligences but very good in solving one particular problem. Concerning this problem, the algorithm is expected to perform like a human, just better. But the very possibility of ‘being better’ is based on a fundamental difference of the algorithmic system and the human in most other regards.

Concerning smart CCTV, the idea is not to model human cognition, but to detect the same event humans would deem noteworthy by evaluating other features. There is a plethora of pattern recognition technologies used for smart CCTV (Hu et al., 2004; Velastin, 2009). One approach just looks at ‘movement trajectories’ of persons (Fuentes and Velastin, 2004; Nguyen et al., 2005; Makris and Ellis, 2002). The paths of people in the video images are reconstructed in three-dimensional space. Then a classification algorithm distinguishes trajectories that could be related to something noteworthy happening from ‘normal’ movement. Such algorithms are believed to provide the same results whether the person has dark or light skin, and do not attribute a gender or cultural background or many other discriminatory factors to the persons under surveillance. Thus such systems promise greater objectivity of the decisions. Of course, this is only partly right. Such a system could for example easily discriminate against persons with disabilities, i.e. ‘nonstandard’ ways of moving. Matzner (2016) discusses these ethical issues in detail. In particular, the paper shows that the presumed independence when the human is meant to have the final responsibility is not given. The judgment is the product of a human-technology ensemble.

The use of movement trajectories is premised on the idea that this single feature suffices as an indicator for the events that an ideal unprejudiced operator would pick out as well. This ideal subject position is enabled by the ensemble of humans and algorithms and cannot be reduced to either side. The algorithm allows to care for the relevant, local details in an objective fashion, making up human ‘weaknesses’ like missing important aspects or imposing overgeneralizing prejudices. But still, the human is meant to judge, not a cold automatic rationality. So the human ‘in the loop’ is not the same like the human ‘on the loop’ or ‘out of the loop’ (DG External Policies, 2012: 6), just with different tasks. To the contrary, the ‘human in the loop’ is a particular subject position enacted by the similarities and differences to algorithms. The human’s opposite, the machinic rationality no longer seems to be a threat as in *Alphaville* because it serves just one particular function – a function where algorithms are enacted on a continuum with humans. The peculiar interplay of likeness and difference thus mobilizes the humanist traits of the first kind of boundary practice, where rationality in the sense of autonomous, objective information processing is foregrounded, in conjunction with the anti-humanism of the second kind, where the human appears as embodied, situated, subjective, dependent. This creates a tension between two perspectives on the human, where the embodied, emotional, located operators appear as lacking compared to a more liberal, autonomous ideal. But this pertains only to one aspect, which is structured by the continuity of humans with machines. Therefore, an algorithm can compensate this lack.

Algorithmic creation: associative and parametric design in architecture and urbanism

Associative or parametric design in architecture are design practices enabled by advancements of software and computation power. Program suites allow to model complex components of a building and their relationships. Rather than drawing the final shape of the building or rooms, as in earlier computer aided design, the architect just specifies a model of internal relationships and parameters that can influence the model. The shape of the building is iteratively generated by running the parameters through the model. This process is guided by ancillary conditions (Rolvink et al., 2010). Such design processes can be used to generate series of uniquely formed parts to build complex, irregular shapes. Current modelling algorithms allow much more complex models that do not just include the form of a building and its parts but also higher level abstractions, like the function of certain rooms, or movement of persons in the building. They also allow to model external factors. As Luciana Parisi writes referring to Michael Hensel and Achim Menges: 'parametric architecture needs to be conceived as a system with a set of finite internal relationships and external forces that inform it and to which it responds' (Parisi, 2013: 104). Algorithms automatically adopt the parameters of the model to these influences and generates a design that optimally suits the external conditions. For example Peter Trummer created the outline for a settlement in the Arizona desert based on the distribution of heat radiation. Highly parametrized and thus malleable housing units are assembled, each with individual parameters set according to its location. The model also takes into account the changes of the heat radiation by the buildings themselves, thus modelling the 'collective behaviour' (Trummer, 2009: 67) of the individual elements. The idea of associative design is thus extended from single buildings to entire settlements. Such models also try to include more social parameters, and not only physical conditions like heat. In a video produced during a workshop with Trummer, the distribution of sunlight and privacy (in terms of lines of sight) are mentioned as explicit parts of a model.[5]

This possibility of associative design is considered as potential to create a new form of architecture. It is decidedly distinguished from the way of building that according to Sjoerd Soeters emerged in the 'early twentieth century': Architects felt like 'gods' that bulldoze the site into a tabula rasa on which they can transform their ideas freely into buildings, creating a 'new order' according to their will (Soeters, 2005: 69). Such an approach to architecture seems to prevail also among the proponents of parametric architecture, who build impressive monuments, often for dubious heads of states in rather authoritarian countries. But others express the potential to displace this centrality of the architect by algorithmic fine-grained attunement of the buildings to the situation. For example, Francois Roche describes his idea for an 'unpredictable, organic urbanism' as a structure that 'develops its own adaptive behavior, based on growth scripts and open algorithms. It is entirely reflexive, responding to human occupation and expression rather than being managed or operated at human will' (Roche, 2009: 42). Similarly, Tom Verebes explicitly wants to 'surpass the mere shaping of a new style, and today's fascination with complex, curvilinear form'. Rather than adhering to these 'deviants of Modernism', he emphasizes the potential to extend the 'invisible informational control systems' and 'augmented cybernetic apparatus[es]' that already manage 'the quotidian fluxes, flows and pulses of the city' into the design (Verebes, 2009: 25).[6] Rather than being the material framework of city life, algorithmic architecture is meant to become reactive to that life. Luciana Parisi discusses projects and ideas that push this potential to its extremes, when architecture is meant to react to 'real-time' inputs based on continuous algorithmic processes – rather than confining them to the design stage (Parisi, 2013: 104). Algorithms then do not only plan the building but also manage and run it.

These discourses discuss associative design as shift from the architect as creator, that is as autonomous liberal subject, towards an algorithm. But this is a limited perspective. As in the case of smart CCTV, decisions and agency are not simply moved but reconfigured. The architect emerges from this reconfiguration as the person responsible for creating the model and choosing the important parameters. Algorithmic systems are better in carrying them out. This shift has a similar structure to the algorithmic appropriation of a situation to match idealized human judgment in smart CCTV.

Again, the boundary is enacted in a combination of continuity and difference. The liberal, autonomous architect is one pole of a tension that creates a lack that can be made up by algorithms: care for local details rather than humanist imposition. The other pole, which is structured by the possibilities of algorithmic design, is a socially responsible and responsive architect that would care for all the detail algorithms allow assessing – much like the ideal operator in the case of smart CCTV. These responsible architects decide to yield important aspects of design to algorithmic generation, thus distancing themselves vividly from the 'deviants of Modernism.' Yet, that kind of responsiveness and responsibility again is enabled by a human-technology ensemble. Its internal boundary creates a continuity between the architect's aims and the optimization and generation features of the algorithm. This boundary mobilizes the anti-humanist strands of the human-machine opposition to posit the algorithm against the patronizing architect: The algorithm cares for the local details and can adapt complex models to it. While this shatters the humanist presumptions of the first position (the godlike architect), the position of the responsible modeler is re-established by creating a continuum of the architect's responsible choices of models and parameters and their algorithmic realization. Here the architect, still in a strong social position compared to security personnel, can actively contribute to enacting this boundary work by emphasizing the decision to yield choices to the algorithm. For example Patrick Schumacher boldly states that 'without parametric life process modelling architecture's task can no longer be adequately addressed' so that 'we will have to reject any architectural design process that does not take advantage of the computational resources as outmoded and substandard' (Schumacher 2015a). But at the same time the continuity of humans and algorithms establishes the architect in a position of liberal autonomy regarding the choice of parameters and the models, the 'social' aspects, which are optimally, and objectively put into practice by the algorithm in continuity with human aims. Consequently, Schumacher (2015b) – and not an algorithm – defines the 'success of the framed life process' i.e. the life within his architecture as the 'ultimate purpose' of design, where 'success' of life can be measured in parameters like 'encounter frequency, interaction diversity, communicative depth'. But this would be impossible without an algorithm that can measure or simulate and act upon these parameters. Summarizing, again the peculiar combination of continuity and difference structures the boundary, enacting a subject position irreducible to either human or algorithm.

The human is dead – long live the algorithm!

Smart CCTV and parametric architecture seem to be an answer to some of the practices 'that have given liberalism a bad name' discussed at the end of the introduction. In particular, the universalist presumption to speak for everyone is addressed in the emphasis that the algorithms provide a detailed, objective analysis of the current situation and thus adapted judgments compared to human overgeneralizations and prejudices or liberal autonomous, god-like creation. This is the anti-liberal strand of the continuity of humans and computation: the human misunderstands his[7] dependence and situatedness, which cybernetic systems and now algorithms aptly grasp.

At the same time, via the influence of the second boundary practice, the algorithmic system gets in a position of objectivity and rationality quite akin to liberal subjectivity, while the human is the embodied, situated, emotional counterpart to this position. I have shown that these two elements contribute to a particular boundary: Since we deal with algorithms, and not the complex cybernetic systems or artificial intelligences, which Hayles discusses, the idea of creating an artificial humanlike being vanishes. Instead the continuity of humans and algorithmic systems is maintained just for one highly specialized tasks. In this regard, algorithmic systems perform task that humans do or would want to do, but better. Yet, this being better is enabled by an algorithmic system that is completely different in most other regards and thus far removed from a better version of the human. To the contrary, it is a way to attend to the local, particular, situation for one specific task. The danger of a machinic rationality taking over thus seems to be averted.

This amounts to a particular transformation of the universalizing traits of liberalism that Hayles discusses. The cases I have discussed enact a socially and ethically competent human being in charge, not a universalizing, overgeneralizing rationality. So the embodied, situated human beings are valued. But this embodiment and situatedness appears as lacking pertaining to one particular task: in my examples detecting suspicious behavior or incorporating external factors in the design. Here human weaknesses or traits seem to prevent the ethically or socially responsible judgment humans are meant to make. This lack, however is transformed by the continuity of humans and algorithms. Irresponsible, prejudiced humans have been a topic long before algorithms. And critique of liberal, autonomous, self-centered authorship is a common motive in philosophy and political theory. But here the problems of these positions are transformed into humans being less able to do what algorithms do. And thus algorithms can make up this lack. The trick is that only in this regard the anti-liberal strands of the boundary are mobilized. Only pertaining to this one task, the human appears as lacking. In this sense, the algorithm does not threaten the human but completes it.

Consequently, while both the human and the algorithm appear in positions of tension with the universalizing presumptions of the liberal, autonomous subject, the human-technology ensemble strengthens rather than weakens this subject position. This structure becomes clearer when looking at authors like Anderson (2008) or Pentland (2012). Their work seems to be a response to the displacement of liberal autonomy by information systems, where humans are just nodes or 'membranes' in a complex flow of information, that already Wiener dealt with. Supported by results from neurology and biology, the authors advocate measuring these complex flows as good as possible and then letting algorithms decide what is best for us. And since our only apparently autonomous decisions cannot be trusted, these algorithmic suggestions are implemented via practices like 'gamification' or 'nudges' (Sunstein, 2014). So algorithms and (more or less mildly) behaviorist practices are meant to help humans in realizing their aims as beings that are 'just nodes in a complex flow of information.' But these are aims they paradoxically seem to have decided upon with the full autonomy of the liberal subject. This paradox appears since the libertarian context of these authors dissimulates the tension that enacts the boundary at work and emphasizes just the attention to local and 'real-time' situations. It ignores that the implicit postulation of a human lack to be ameliorated by information technology itself is the product of this particular enactment of the boundary.

This boundary structure can also be found in other applications. Self-tracking is a good example that is often conceived as a means to realize one's potential, in the vein of the aforementionend authors. This idea of potential is the positive reformulation of the implicit lack brought about by human-algorithmic boundaries (where there is lack, one can get better).

This strange alliance of anti-liberal tropes and a highly libertarian atmosphere is a good example for the strengthening of liberal subject positions by human-technology ensembles. The posthumanist critiques of liberalism often focus on decentering the rational, autonomous subject by emphasizing its dependence and exposedness to non-human actants – in this case information technology. Yet, my analysis shows that this emphasis alone does not suffice to challenge the liberal subject position. In fact, this challenge can easily be turned around since technology is not only enacted as the opposite of this humanist subject, but also as its continuation – in particular 'intelligent' information technologies. When this continuity comes into play concerning one small task – and not a transhumanist project – the dependence, which would decenter the human, turns into a lack that the algorithm can compensate. The non-human agents then contribute to a human-technology ensemble that *in toto* strengthens rather than weakens the liberal, autonomous subject position. And faithful to its tradition, this position ignores that it is not a potential for all humans, but only for some: those that can afford the technology, those that are deemed worthy by the economic efficiency that structures both cases I have discussed, those that can harness their use of new technologies as advancement – like the responsible architects do.

Discussions what algorithms can do and whether they challenge human freedoms might distract from the fact that within these discussions, a tension between the posthuman or anti-liberal positions and their (often implicit) liberal, humanist counterparts are at issue. Donna Haraway introduces her famous celebration of hybrids, the *Cyborg Manifesto*, as 'argument for pleasure in the confusion of boundaries and for responsibility in their construction' (1991: 150). Maybe the first of these two aspects has been too much in the focus in attempts to decenter the liberal subject. Although authors like Latour (1993) have distanced themselves from projects of disclosing ideologies or false consciousness, many advocates of networks, hybrids, assemblages may have relied too much on just showing how the liberal subject is actually a product thereof, rather than critically aiming at different productions. In this sense, the advocates of the algorithmic technologies I have discussed may indeed have learned a lesson. The human is not the autonomous, rational subject that liberal and humanist discourses talk about. But with the help of algorithms, as human-technology ensembles, the liberal subject that the human alone never has been, finally seems possible. With the detailed analyses I presented here we can see how human-algorithmic ensembles strengthen liberal, autonomous subject positions, rather than decentering them.

This entails that the focus on *algorithms* as actors, as opposites or competitors of the human facilitates the enacting of this subject position. Many of the discussions what algorithms can or cannot do thus implicitly legitimize the underlying liberal subject positions – including the 'practices that have given liberalism a bad name'. This also pertains to recent warnings for careful research on artificial intelligence (Cellan-Jones, 2014). Again, the specter of some hyperintelligent, super-human is foregrounded as the threat posed by artificial intelligence – rather than seeing how algorithmically driven systems already shift subject positions and power relations. Bringing the material, discoursive, social, mathematical, etc. preconditions into the picture is an important step but does not suffice. The important issue is not only descriptive accuracy what algorithms are or can do, but which subject positions – and that is power positions – are created and legitimized. A critical inquiry concerning algorithms thus should include inquiring potential new boundaries and subject positions, taking 'pleasure in the confusion of boundaries' but maybe even more importantly, 'responsibility in their construction.'

---

[1] Transcendence. Dir. Jack Paglen; Alcon Entertainment, DMG Entertainment and Straight Up Films (2014).
[2] Her. Dir. Spike Jonze, Annapurna Pictures (2013).
[3] Alphaville, une étrange aventure de Lemmy Caution. Dir. Jean-Luc Goddard, Athos Films and Chaumiane (1965).
[4] Ex machina. Dir. Alex Garland, DNA Films and Film4 (2015).
[5] https://www.youtube.com/watch?v=EhjUli4cYEg (last access 29.09.2015).
[6] That a lot of these control systems are parts of huge apparatuses of surveillance and governance (Lyon, 2005) does not seem to worry these authors. Albeit, one could easily conceive this outlook as an update of the alliance between architecture and governance and policing to societies of digital control.
[7] Here the male gender is on purpose.

References

Amoore, L. (2011) 'Data Derivatives: On the Emergence of a Security Risk Calculus for Our Times', *Theory, Culture & Society* 28(6): 24–43.

Anderson, C. (2008) 'The End of Theory: The Data Deluge Makes the Scientific Method Obsolete', *Wired* 16(7).

Anderson, M.L. (2003) 'Embodied Cognition: A field guide', *Artificial Intelligence* 149(1): 91–130.

Barad, K. (2007) *Meeting the Universe Halfway*. Durham: Duke University Press.

Berry, David M (2011) *The philosophy of software: code and mediation in the digital age*. Basingstoke: Palgrave Macmillan.

Blanchette, J. (2011) 'A material history of bits', *Journal of the American Society for Information Science and Technology* 62(6): 1042–1057.

Blanchette, J. (2012) 'Computing as if infrastructure mattered', *Communications of the ACM* 55(10): 32–34.

Blass, A. and Gurevich, Y. (2003) 'Algorithms: A quest for absolute definitions', *Bulletin of the EATCS* 81: 195–225.

Braidotti, R. (2013) *The Posthuman.* London: Polity.

Brey, P. (2005) 'The Epistemology and Ontology of Human-Computer Interaction', *Minds Mach* 15(3-4): 383–398.

Cellan-Jones, R. (2014) 'Stephen Hawking warns artificial intelligence could end mankind' URL (consulted April 2016): http://www.bbc.com/news/technology-30290540.

Chun, W. H. K. (2008) 'On "sourcery," or code as fetish', *Configurations* 16(3): 299–324.

Citron, D. K. and Pasquale, F. A. (2014) 'The scored society: due process for automated predictions', *Washington Law Review*, 89, 1-33.

Daston, L. (1988) *Classical probability in the Enlightenment*. Princeton University Press.

Dee, H. M. and Velastin, S. A. (2008) 'How close are we to solving the problem of automated visual surveillance?' *Machine Vision and Applications* 19(5-6): 329–343.

Diakopoulos, N. (2013) 'Algorithmic Accountability Reporting: On the Investigation of Black Boxes.' A Tow/Knight Brief. New York: Columbia Journalism School, Tow Center for Digital Journalism. URL (consulted April 2016) http://www.nickdiakopoulos.com/wp-content/uploads/2011/07/Algorithmic-Accountability-Reporting_final.pdf

Directorate-General for External Policies of the European Union (2012) 'Human rights implications of the usage of drones and unmanned robots in warfare', URL (consulted April 2016): http://www.europarl.europa.eu/committees/en/studies.html

Drucker, J. (2013) ‘Performative Materiality and Theoretical Approaches to Interface’, *Digital Humanities Quarterly* 7(1).
Fuentes, L. and Velastin, S. (2004) ‘Tracking-based event detection for CCTV systems’, *Pattern Analysis and Applications* 7(4): 356–364.
Gillespie, T. (2014) The Relevance of Algorithms, pp. 167-194 in Gillespie, T. Boczkowski, P and Foot, K. (ed.) *Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA: MIT Press.
Gillespie, T. (2011) Can an algorithm be wrong? Twitter Trends, the specter of censorship, and our faith in the algorithms around us. URL (consulted April 2016): http://culturedigitally.org/2011/10/can-an-algorithm-be-wrong/.
Haraway, D. (1991) ‘A Cyborg Manifesto - Science, Technology, and Socialist-Feminism in the Late Twentieth Century’, pp. 149-182 in *Simians, Cyborgs and Women: The Reinvention of Nature*. New York and London: Routledge.
Haraway, D. J. (1997). *Modest_Witness@Second_Millennium. FemaleMan©_Meets_ OncoMouse™: Feminism and Technoscience.* New York and London: Routledge.
Hayles, N. K. (1999) *How We Became Posthuman*. Chicago: University of Chicago Press.
Hayles, N. K. (2010) ‘How We Became Posthuman: Ten Years On. An Interview with N. Katherine Hayles’, *Paragraph* 33(3): 318–330.
Hu, W., Tan, T., Wang, L. and Maybank, S. (2004) ‘A survey on visual surveillance of object motion and behaviors’, *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* 34(3): 334–352.
Introna, Lucas D. (2016) ‘Algorithms, Governance, and Governmentality: On Governing Academic Writing’, *Science, Technology & Human Values* 41(1): 17-49.
Kitchin, R. (2014) ‘Big Data, new epistemologies and paradigm shifts’, *Big Data & Society*, 1(1), 1-12.
Kitchin, R. and Dodge, M. (2011) *Code/Space: Software and Everyday Life.* Cambridge, MA: MIT Press.
Kleene, S. C. (1952) *Introduction to Metamathematics*. Amsterdam: North Holland.
Knuth, D. E. (1973) *The Art of Computer Programming, Vol. 1*. Reading: Addison-Wesley.
Latour, Bruno (1993) *We have never been modern*. Cambridge, MA: Harvard University Press.
Lyon, D. (2005) *Surveillance as Social Sorting*. London: Routledge.
Neyland, Daniel (2015) ‘On Organizing Algorithms’, *Theory, Culture & Society*, 32(1), 119-132.
Mackenzie, A. (2015) ‘The production of prediction: What does machine learning want?’ *European Journal of Cultural Studies* 18(4-5): 429–445.
Makris, D. and Ellis, T. (2002) ‘Path detection in video surveillance’, *Image and Vision Computing* 20(12): 895–903.
Matzner, T. (2016) ‘The model gap: cognitive systems in security applications and their ethical implications’, *AI & Society* 31 (1), 95-102.
Nguyen, N. T., Phung, D. Q., Venkatesh, S. and Bui, H. (2005) ‘Learning and detecting activities from movement trajectories using the hierarchical hidden Markov model’, pp. 955–960 in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.*
Parisi, L. (2013) *Contagious architecture: computation, aesthetics, and space.* Cambridge: MIT Press.
Pentland, A. (2012) *Reinventing Society in the Wake of Big Data*. URL (consulted September 2014): http://www.edge.org/conversation/reinventing-society-in-the-wake-of-big-data.
Roche, F. (2009) ‘I’ve Heard About…(A Flat, Fat, Growing Urban Experiment): Extract of Neighbourhood Protocols’, *Architectural Design* 79(4): 40–45.
Rolvink, A., van de Straat, R. and Coenders, J. (2010) ‘Parametric structural design and beyond’, *International Journal of Architectural Computing* 8(3): 319–336.
Schumacher, P. (2015a) Fluid Totality – The dream of inhabiting a nature-like built environment. URL (consulted April 2016): http://www.patrikschumacher.com/Texts/Fluid_Totality.html.

Schumacher, P. (2015b) Parametricism with Social Parameters. URL (consulted April 2016): http://www.patrikschumacher.com/Texts/Parametricism%20with%20Social%20Parameters.html
Simons, D. J. and Chabris, C. F. (1999) 'Gorillas in our midst: Sustained inattentional blindness for dynamic events', *Perception-London* 28(9): 1059–1074.
Soeters, S. (2005) 'On being a humble architect', pp.69-74 in Ray N (ed.) *Architecture and its ethical dilemmas*. New York: Taylor & Francis.
Sunstein, C. R. (2014) *Why nudge?: the politics of libertarian paternalism*: Yale University Press.
Totaro, P., and Ninno, D. (2014) 'The Concept of Algorithm as an Interpretative Key of Modern Rationality.' *Theory, Culture & Society* 31(4): 29-49.
Trummer, P. (2009) 'Morphogenetic Urbanism', *Architectural Design* 79(4): 64–67.
Turing, A. M. (1937) 'On computable numbers, with an application to the Entscheidungsproblem', *Proceedings of the London Mathematical Society* 2(42): 230-265.
van der Ploeg, I. (2005) Biometrics and the body as information: Normative issues of the socio-technical coding of the body, pp. 57-73 in Lyon D (ed.) *Surveillance as Social Sorting*. London: Routledge.
Velastin, S. A. (2009) 'CCTV Video Analytics: Recent Advances and Limitations', pp. 22–34 in *Proceedings of the 1st International Visual Informatics Conference on Visual Informatics: Bridging Research and Practice*. Berlin, Heidelberg: Springer.
Verebes, T. (2009) 'Experiments in Associative Urbanism', *Architectural Design* 79(4): 24–33.
Williams, K. S. and Johnstone, C. (2000) 'The politics of the selective gaze: closed circuit television and the policing of public space', *Crime, Law and Social Change* 34(2): 183–210.
Wolfram, S. (1984) 'Cellular automata as models of complexity', *Nature* 311(5985), 419-424.