



Infosphere, Datafication, and Decision-Making Processes in the AI Era

Andrea Lavazza¹ · Mirko Farina^{2,3}

Accepted: 5 April 2023

© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract

A recent interpretation of artificial intelligence (AI) (Floridi 2013, 2022) suggests that the implementation of AI demands the investigation of the binding conditions that make it possible to build and integrate artifacts into our lived world. Such artifacts can successfully interact with the world because our environment has been designed to be compatible with intelligent machines (such as robots). As the use of AI becomes ubiquitous in society, possibly leading to the formation of increasingly intelligent bio-technological unions, there will likely be a coexistence of a plethora of micro-environments wrapped and tailored around humans and basic robots. The key element of this pervasive process will be the capacity to integrate biological realms in an infosphere suitable for the implementation of AI technologies. This process will require extensive datafication. This is because data is the basis of the logical-mathematical codes and models that drive and guide AI. This process will have huge consequences on workplaces, on workers, as well as on the decision-making processes required for the functioning of future societies. In this paper we will offer a comprehensive reflection on the moral and social implications of datafication as well as a set of considerations about its desirability, which will be informed by the following insights: (1) full protection of privacy may become structurally impossible, thus leading to undesirable forms of political and social control; (2) worker's freedom may be reduced; (3) human creativity, imagination, and even divergence from AI logic might be channeled and possibly discouraged; (4) there will likely be a push towards efficiency and instrumental reason, which will become preeminent in production lines as well as in society.

Keywords Infosphere · Datafication · Decision-making · AI · Ethics

1 Introduction: Infosphere and Datafication

A recent interpretation of AI developments proposes to consider AI as a form of acting that does not have to be intelligent to be successful (Floridi, 2013, 2022). The basic idea is to return to how the problem of intelligence was framed

by the initiators of contemporary cognitive science (McCarthy et al. 2006). According to Floridi, it is sufficient to have recourse to a counterfactual, which concerns human behaviour. In this sense, the problem of artificial intelligence is only that of making a machine act in ways that would be called intelligent if a human being behaved in the same way. Thus, there is no issue of comparison between human intelligence and machine intelligence. The only relevant issue is to perform a task successfully, such that the result is as good or better than human intelligence would be able to achieve. How this happens is not the central issue (although it may have important consequences); the outcome is.

This approach to AI is called engineering or reproductive. It aims to reproduce the results or successful outcome of our intelligent behaviour by nonbiological means. In contrast, the cognitive or productive approach to AI aims to produce the nonbiological equivalent of our intelligence; that is, the source of the behaviour that the engineering approach aims to reproduce (cf. Floridi 2011a, b).

Mirko Farina

m.farina@innopolis.ru

Andrea Lavazza

lavazza67@gmail.com

¹ Centro Universitario Internazionale, Via Antonio Garbasso 42, Arezzo 52100, AR, Italy

² Human Machine Interaction Lab, Faculty of Humanities and Social Sciences, Republic of Tatarstan, Innopolis University, Universitetskaya St, 1, Innopolis 420500, Russia

³ Lab of Industrializing Software Production (LIPS), Faculty of Computer Science and Engineering, Republic of Tatarstan, Universitetskaya St, 1, Innopolis 420500, Russia

The reproductive approach has achieved astounding successes very quickly and promises to continue to advance exponentially (think about the development of mRNA vaccines against COVID-19 (Pizza et al. 2021), where being able -thanks to AI tools- to do reprogramming (as fast as possible and in a way that is as coordinated as possible) helped managing the deluge of data associated with the project¹. In so many areas, reproductive AI is better and tends to replace human intelligence because it is faster, more reliable, and more consistent in its results. The absolute reliability of AI in standardized tasks is probably the main difference with a human operator. While the latter may manifest greater degrees of freedom in task execution (something to do with the cognitive and productive aspects of AI), automated systems guarantee unambiguous, infinitely repeatable performance without fluctuations.

This is also a cultural feature that cannot be overlooked when considering the labour market and industrial production. Indeed, one can identify a characteristic sought by both the supply side and the demand side; namely the desire for a product that is “perfect” insofar as it is not limited or influenced by human “imperfections”. The delegation to a “dumb”-as Floridi calls it-but exceptionally effective AI allows us to make our lives much easier and less tiring. AI as a reservoir of capabilities can therefore tackle any number of problems and tasks for which human intelligence characteristics of understanding, awareness, sensitivity, semantics, and meaning are not needed. And this happens, as proposed by Floridi (2013, 2014, 2022) among others, as the world adapts to reproductive AI and not vice versa.

Industrial automation follows this paradigm. The introduction of robots or devices that carry out the production and distribution processes with reduced human intervention or diminishing participation is done by circumscribing the work environment to the limited capabilities of simple machines. We don’t try to build a humanoid robot to wash clothes in a bathtub but build a microenvironment (such as a washing machine) that takes advantage of available technology. The same happens with automated ironing. This changes not only the way people work towards the realization of these activities, but also the products for which the services are designed. We are talking here about technologies that are not cutting-edge, where AI plays a limited role.

Consider, however, other procedures, such as house cleaning. Robot vacuum cleaners take advantage of AI to move with increase effectiveness in complex environments. However, it is clear that it will soon be the design of homes that will adapt to automated service systems, especially with

the needs of the elderly in mind, if robotic assistants become more prevalent for lonely people.

The self-driving car may be one example among the high-tech ones, where engineering AI is the absolute protagonist (Bonnefon 2021). The self-driving car does not start out as a classic self-driving car that is adaptable to different road locations and can, if need be, travel on unpaved terrain or in adverse environmental conditions, such as a blackout of lighting and electronic signage. The self-driving car comes with specific requirements due to AI technology that allows the vehicle to move without a human driver. It must move in an environment that allows it to have all the feedback necessary for the efficient execution of its task, which is to move from point A to point B with maximum safety and comfort of the passengers and all who may be in its path. This can be accomplished by engineering the roads, making them suitable for the self-driving car (Birdsall 2014). It is not the car that has to adapt to the environment, but it is the environment that is wrapped-around a tool that we find particularly useful in terms of saving effort, time, and traffic accidents (Borenstein et al. 2019). Paradoxically, at an early stage, self-driving cars will have a narrow range of available destinations and thus condition the mobility of those who want to rely on them. For instance, robotaxis can only circulate on a few streets in San Francisco (cf. Heaven 2022) or in very small cities (such as Innopolis).

In general, wrapping the environment in an infosphere has become an increasingly common practice to exploit the potential of AI, where “the infosphere is the whole system of services and documents, encoded in any semiotic and physical media, whose contents include any sort of data, information and knowledge (...) with no limitations either in size, typology, or logical structure. Hence it ranges from alphanumeric texts (i.e., texts, including letters, numbers, and diacritic symbols) and multimedia products to statistical data, from films and hypertexts to whole text-banks and collections of pictures, from mathematical formulae to sounds and videoclips” (Floridi 1999).

Connected to the infosphere is the onlife dimension, i.e., the activity that everyone performs while being connected to digital devices, which are also embedded in the wrapping-around logic we referred to above. Environments are changing, so that artificial agents-robots, bots, algorithms-can move with greater ease than humans can now do. In highly digitally wrapped environments, all relevant data are collected (or at least potentially collected) and analysed without the need for other interventions. Thus, decisions and actions can be made automatically by applications and actuators.

In this context consider the process of datafication, which is illustrative of many of the ideas discussed above. Datafication, according to Mayer Schoenberger and Cukier

¹ The reader can find more details about this at: https://podcasts.apple.com/us/podcast/i-was-there-when-ai-helped-create-a-vaccine/id1523584878?i=1000577165082&mc_cid=9a3b4e96f5&mc_eid=5282de56e0, Last accessed April 2023.

(2013a, b) is the transformation of social action into online quantified data; a procedure that allows for real-time tracking and predictive analysis of consumers' behaviors. Simply stated, datafication is all about accessing -with the help of AI tools - previously inaccessible processes or activities and turning them into data, that can be subsequently monitored, tracked, analyzed and optimized, or even sold (Cukier and Mayer-Schoenberger 2013). To be sure, the exploitation of Big Data can unlock significant value in areas such as decision making, customer experience, market demand predictions, product and market development and operational efficiency (Yin and Kaynac 2015) and many of the technologies we use in our daily life have enabled different ways of 'datafying' our basic activities and behaviors (Da Bormida 2021).

Social networks (such as Facebook or Instagram) notoriously collect and monitor data information to market products and services with the intent to produce recommendations to potential buyers (Chamorro-Premuzic et al. 2017). Yet, datafication is a much more pervasive phenomenon than the naïve eye may prima facie meet, as it is actively pursued (with different goals and aims) by many industries (Pybus and Coté 2021), for example:

- by insurance companies, where the data gathered is used to update risk profile development and business models;
- by banks, to establish the trustworthiness of a certain individual requiring -for example- a loan;
- by human resources and hiring managers at various level, which use datafication to identify risk-taking profiles or even to spot potential personality issues;
- by governments and institutions, where datafication and digitalization are often pursued with the intent of minimizing bureaucracy and optimizing transparency in both decision making and resource allocation;
- (in general), by investors worldwide to boost business opportunities, credentials, and productivity. For example, very successful companies (such as Netflix, Amazon, Uber, Fitbit) typically merge the resourcefulness of big data with the power of AI to offer their users products that are smart and reliable.

In short, one can argue that datafication -especially if pursued in an infosphere- can make our lives smoother and -in doing so- fundamentally change our societies, how people interact between each other and with their institutions, and probably even transform people's understanding of the concept of community as a whole (Skenderija, 2008).

Nevertheless, in the face of these positive effects any data-driven endeavor that takes place in an infosphere must also be considered (and therefore properly assessed) against the backdrop of complex and multidimensional issues or

challenges that it may contribute to form, concerning -for instance- decision-making processes, social solidarity, privacy, security, the management of public goods, of civil liberties, or even sovereignty (Da Bormida 2021).

For example, in the health care sector, concerns about the datafication of the infosphere relate to the difficulty of respecting ethical boundaries relating to sensitive data (e.g., Ruckenstein and Schüll 2017). Datafication, it has been argued, has the potential to erode goal orientation and the room for professional judgement (Hoeyer and Wadmann 2020), favoring varieties of neoliberal subjectification (Fotopoulos and O'Riordan 2016; Foucault 1991) in the form of tools that may accelerate the withdrawal of the welfare state from citizens' lives, which can eventually turn health care into self-care (Ajana 2017).

In the education sector, the major risk involved is that students may feel constantly under 'liquid surveillance' (Bauman and Lyon 2013; Zuboff 2019), due to the continuous collection and processing of their data on all levels of their learning trajectory in the educational system (from the classroom to the school, from the region to the state and internationally [Jarke and Breiter 2019]). This – it has been observed- can potentially lead to a reduction of their creativity and/or in higher levels of stress (Williamson et al. 2020).

Thus, while wrapping up environments to harness the potential of AI represents a good way to improve humans' condition, the future of our lives is (and will be even more) marked by datafication, which may actively modify our environments in the attempt to achieve more effectiveness and efficiency. The modification of work processes pursued within the infosphere of an increasingly datafied society has several consequences. While many researchers have investigated the consequences of datafication in separate fields (e.g., Da Bormida 2021), not much work has been done -so far- to bring all these insights together in one research paper. This is what we propose to do in our contribution.

Specifically, we show that datafication in a rich infosphere may determine that: (a) the full protection of privacy may become structurally impossible, thus leading to undesirable forms of political and social control; (b) worker's degrees of freedom and security may be reduced; (c) creativity, imagination, and even divergence from AI logic might be channeled and possibly discouraged; (d) there will likely be a push towards efficiency and instrumental reason, which will become preeminent in production lines as well as in society. All this encourages reflections on the ways in which digital technologies may foster or hinder decision-making processes in future societies and on how increasingly automatized algorithms, based on machine learning, may gradually take over certain roles that were previously uniquely attributed to humans.

This does not necessarily mean that the development of AI to improve working conditions should be resisted; rather, we should reflect on how to better organise the process to achieve social and moral good. The first concern of ethics in the face of the advance of AI is with workers and their condition. The goal is therefore to identify the risks that individuals and society at large may face and to find regulatory remedies to those risks. In the next four sections, we will look at areas where the spread of AI in workplaces and processes may require conceptual clarification and both ethical and legislative regulation.

2 Privacy Issues

Several researchers working on datafication (e.g., Van Dijck 2014) argue that surveillance is ‘too optically freighted and centrally organized a phenomenon to adequately characterize the networked, continuous tracking of digital information processing and algorithmic analysis’ (Ruckenstein and Schüll 2017, p.264), that occurs in the world in which we nowadays live. On these grounds, such researchers propose to replace the term ‘surveillance’ with the term ‘dataveillance’ (Gitelman 2013; Ruppert 2011), by which they mean that the act of surveillance in today’s world does not take place directly from the above, but rather becomes distributed across multiple parties and several domains (covering much of our activities and potentially spanning from business to education, from medicine to justice, from governance to management).

These researchers (e.g., McQuillan 2016) also notice a different telos (or end goal) between surveillance and dataveillance. Where the end goal of surveillance might be defined as the ability to constantly ‘see’ something or someone; the telos of dataveillance is rather concerned with the capability of continuously tracking information across multiple domains to capture emergent patterns capable of predicting people’s behaviors (not only of observing it). Yet, algorithms and tracking AI tools are not only used to detect and predict one’s behavior but also to shape and actively modify it (Beer 2009; Mackenzie 2005).

For example, the data that users generate might be gathered and processed to give a digital feedback capable of indirectly modulating and orienting someone’s action, in a way that subtly departs from direct panoptic forms of discipline but could be argued to be even more effective. An illustration of this claim is the growing usage of wellness programs in corporate settings (Till 2017). Such programs typically encourage employees—through incentives or rather penalties—to engage in self-tracking activities, with the intent of gathering data that employers (in various forms and at various levels) can then analyze, by using proprietary

algorithms (Christophersen et al. 2015). As Kennedy et al. (2015, p. 1) brilliantly put it: ‘the advent of big data brings with it new and opaque regimes of population management, control, discrimination and exclusion’, something very much akin to what Foucault (1997) called biopolitics; a pervasive mode of power that attempts to understand, control, influence and even regulate the vital characteristics of any given population (Farina and Lavazza 2021b). In agreement with Lupton (2016), we believe that we are now entering an era in which biopolitics may be enforced through datafication; that is, through the joint combination of extensive datasets of digital information gathered synchronously across multiple domains. All this raises crucial issues surrounding the privacy of individuals as well as their basic civil liberties (such as freedom of movement and freedom of association) that are now –it seems to us– more than ever under threat (Farina and Lavazza 2021b; Pietrini et al. 2022; Lavazza & Farina 2021).

Consider the following example as a paradigmatic illustration of this claim. It involves the collection of biometric data through face recognition algorithms based on machine learning (Gray and Henderson 2017; Ball et al. 2012). This is just an example of a more general trend (involving the application of biometrics in society). We note that the rolling out of this technology is taking place as we write this paper in many countries, especially in those in which there is a rich infosphere that supports widespread technological advancements (such as the development of 5G).

Biometrics can be defined as ‘the science of automatic identification or identity verification of individuals using [unique] physiological or behavioral characteristics’ (Vacca 2007, p.589). Roughly speaking, biometric systems can be divided into two main categories: hard biometrics and soft biometrics. Hard biometrics include traditional biometric identifiers (such as faces, iris scans, DNA markers, and fingerprints) that are normally used for identity verification technologies (Benziane and Benyettou 2011). Soft biometrics are instead parameters (such as gender, ethnicity, age, height, weight, voice accent, birthmarks etc) that can complement hard biometrics and be used to increase the precision or the accuracy of the recognition system (Nixon et al. 2015). Soft biometrics typically provides information about a person, without –on its own– necessarily providing sufficient evidence to precisely determine the identity of that person.

The process of biometric identification is quite complicated and can be summarized in four basic steps (Hu 2017), which include: (1) *Enrollment* (biometrics data are gathered from the individual); (2) *Recognition* (a template of the individual’s identity is created on an artificial system for monitoring purposes); (3) *Comparison* (future biometrics data are gathered from individuals); and (4) *Decision* (a

match is found or not found among the data collected based on specific algorithms that cross-check all biometrics data obtained on the individual). We note that biometric database screening technology is increasingly employed in this fourth step, as it is believed to remove the human element from the matching process, thereby maximizing objectivity and efficacy in decision-making (Ellerbrok 2011).

Biometric technology is also increasingly considered as an effective tool for dealing with security matters (such as terrorism prevention). Because of this, the last decade has seen a very rapid development of biometric technologies (Alsaadi 2021). ‘Biometric datavillance programs’, as we may call them, are proliferating under preemptive strategies to combatting crime and terrorism and to ensure homeland as well as international security. We shall note that the U.S. Department of Defense (DoD) has called such approaches -perhaps in a Freudian slip - ‘population management’², which suggests that their potential applications may well stretch -to put it mildly- to much wider realms, quite possibly along the lines envisaged by Foucault (1997)³.

Anyhow, major recent trends in biometrics typically focus on individuating behavioral kind or towards the development of ‘multimodal biometrics’ (Ryu et al. 2021), a procedure which involve the combination of sensor and computing capabilities endowed with enhanced connectivity with the intent to apply such technologies in a broad variety of sectors and for a broad variety of purposes, far beyond law enforcement or prevention of crimes (Hu 2017). For example, latest breakthroughs in the field include the development of sensors that can capture new types of bio-signals (such as heart beats and brain waves via -for instance- EEG or ECG), or brain-computing-interfaces (BCI).

Such interfaces are reported to be able to measure neuro activity and translate it into machine-readable inputs (Anumanchipalli et al. 2019), which suggests that these devices could -in the future- allow for the detection of thoughts, possibly opening to the possibility of influencing operations of the human brain. We won’t focus on such technologies on this paper, as they are mostly covered by state secrets (and are currently under development); however, we would like to spend the remainder of this section on analyzing the case of face recognition technology through machine learning algorithms, which is equally significant and perhaps is the one that poses -at this stage at least, especially given its widespread application in society- the most significant ethical and social challenges.

Humans are very good at recognizing fellows based on facial appearance. Naturally then, face can be considered

as an ideal trait to use for automated biometric recognition. Face recognition systems typically utilize the spatial relationship among the locations of facial features (such as eyes, nose, lips, chin, and the global appearance of a face [Jain 2007]) in conjunction with rapidly developing artificial intelligence (AI) technologies, to provide information that can be used for security and law enforcement purposes. See Ali et al. (2021) and Boutros et al. (2022) for surveys of recent face recognition technologies.

For example, western countries (such as United Kingdom, United States, and Australia), being at the forefront of the development of comprehensive surveillance systems, increasingly use such technologies for security purposes (without getting into unnecessary technicalities, anyone walking around London can easily get a feeling of that).

The expanding use of this technology therefore raises pressing ethical and social concerns regarding its adoption in society. ‘Central to the ethical, legal and policy issues is the tension that exists between the legitimate collection of biometric information for law enforcement, national/international security, and government service provision, on the one hand; and the rights to privacy and autonomy for individuals on the other’ (Smith and Miller 2022, p.168). Descending from this point there are also issues concerning potential violations of individuals’ privacy in search of wrongdoings that can lead to imbalance between a state and its citizenry and that need to be carefully evaluated.

In modern societies, it is normally agreed that the state has no right to engage in selective monitoring of any citizen, unless that citizen raised strong suspicions of unlawful behaviors. Yet, the development of facial identification technology invites the active monitoring and even the full-scale mapping of law-abiding citizens; in essence, the pervasive wrapping of technology around innocent civilians, which may contribute to undermine the basic universal right of not being investigated selectively (Gstrein and Beaulieu 2022).

Of course, face recognition technology is also used for good things. For example, it is widely deployed in airports, where it has contributed to speed up the processing on incoming passengers by customs authorities. Legislation to facilitate the usage of facial recognition programs capable of integrating pictures from passports and various forms of IDs (such as drivers licenses) into a national database, which can then be consulted by law enforcement and other government agencies are being introduced in several countries across the globe; however, the average reader is probably less aware that such technology is also being actively rolled out in many countries, especially in connection with the development of 5G networks.

5G networks, which possess extremely high computational power combined with the huge storage capability of modern clouds (we are talking about zetta possibly

² <https://publicintelligence.net/identity-dominance/>, Last accessed April 2023.

³ The Snowden disclosures can be considered as a paradigmatic illustration of this claim (Lyon 2014).

yottabytes of images and even videos), represent the ideal companion for this facial recognition technology in as much as they allow to fully exploit its potentials in richly datafied and infospheric environments. In brief, current face recognition technologies allow to store huge amounts of personal data coming from multiple domains and timespans, to reliably access them at will at any point in time, with fast algorithms specifically designed to selectively checking all the information gathered for ‘desired’ purposes.

Yet, facial recognition programs are -to date, at least- quite vulnerable to deepfake-based attacks (see Ramachandra and Busch 2017, for a helpful review), for example- with static facial images⁴, which raise concerns about the security as well as the effective trustability of those data. In addition, facial recognition technologies might be combined with AI tools preprogrammed for spotting specific emotions (e.g., anger) to target minorities (e.g., prone to rebellion) based on ethnicity (so on automated analyses of morphological traits); hence, they could be massively deployed to discriminate and even oppress -given the pervasivity of such systems in modern infospheric societies- certain strata of any given population (those that -for instance- do not adhere to a state religion due to different cultural backgrounds).

Furthermore, given the storage capabilities of modern clouds, which are set to increase dramatically over the next decades, who could guarantee that the biometric data stored in archives now, through the extensive process of datafication, wouldn’t become compromising -say- 30 years from now, when certain moral values or virtuous might have changed, partly or entirely? Who could then assure that law-abiding citizens couldn’t be prosecuted in 30 or 40 years for behaviors, words, or actions that are completely acceptable now but may not be deemed as ‘convenient’ in the future if a track record of their actions associated with their morphological traits is permanently stored (and readily accessible) somewhere? Given current trends on cancel culture and the corresponding emergence of ‘dataveillance’, this possibility shouldn’t be too hastily ruled out.

These are very crucial issues underlying the usage of facial recognition in biometrics mapping that promise to bear a significant ethical and legal impact on the future of our societies. Having briefly reviewed them, we now look at another application of datafication in rich infospheric environs, which include industrial automation.

3 Freedom Issues

Increasingly datafied working environments are geared toward efficiency and, therefore, in general terms toward reducing worker discretion. In this sense, a certain loss of worker’s ‘freedom’ is inherent – and perhaps even acceptable – in any process, not only of automation but also of standardization and compartmentalization of resources and procedures. The shift from the craftsman performing the whole process of pin production to the division of labour among workers performing different tasks was famously described by Adam Smith in the 18th century. We should now consider the peculiarity of working in an environment that extensively relies on datafication and is richly wired and interconnected (infospheric) across multiple domains and dimensions.

In such an environment, the human being must be a facilitator of processes that automated systems are not yet able to do or will never be able to do. For example, in warehouses this happens with the substantial homologation of workers to the procedures, rhythms, and forms of control and evaluation that have been introduced for processes carried out entirely by industrial robots (Delfanti 2021; Engstrom and Jebari 2022). It is not intended here to make a social and political critique of this kind of evolution of the work environment decoupled from technical considerations referring to productivity gains that translate into concrete benefits for consumers in terms of product availability and low costs. In our societies, all workers are also consumers, and this cannot be underestimated.

However, the more we make the working environment wrapped around robots, the more the risk grows that even human employees will be totally absorbed in this new production procedure, which may have strong repercussions for workers. This could lead to new forms of exploitation, as some are afraid of. Yet, it is not necessarily the case that this will happen. In any circumstance though, the logic of quantification and automation entails a modification of the worker’s spaces of freedom. Indeed, it should be emphasized that two of the basic criteria of AI-based approaches are predictability and certainty. These criteria are structurally opposed to the classical idea of freedom, which is understood as the possibility of choosing from time to time between alternative courses of action based on reason (Lavazza and Inglesi 2015).

There are several areas in which workers’ freedom might be diminished because of the widespread implementation of AI tools and datafication in society. Personnel selection is one of such areas, where the hiring process is progressively being managed by algorithms capable of evaluating candidates in accordance to predefined criteria, which are set against the perceived compatibility of a subject for a

⁴ <http://pinktentacle.com/2008/06/magazine-photos-fool-age-verification-cameras/>, Last accessed April 2023.

specific task. There is already a rather large literature on the possible biases introduced by such programs (Tippins et al., 2021; Goretzko and Israel 2021). These biases depend on how the programs were designed and on the type of data on which they were fed and trained. Typical examples of bias introduced by personnel selection programs trained on time series or previous informal criteria adopted by companies involve decisions unfavourable to women, ethnic minorities, or social groups that have been historically disadvantaged or excluded.

To be sure, discrimination in the workplace has always existed, and power relations between firms and individual workers have always been highly unbalanced and asymmetrical. A new sensibility in recent decades, however, has brought new attention to the issue and has made it possible to reduce systematic bias in selection and various types of abuses (Woods et al. 2020). Yet, the introduction of algorithms that are considered more efficient and unbiased may—if not properly supervised—risk introducing the very same systematic discrimination that we strove to fight in recent decades (Farina et al. 2022a, b; Bakare et al. 2022; Bugayenko et al. 2023). In addition, this new sort of potential systemic discrimination may be far less detectable (as based on mathematical data, which are difficult to interpret for the non-experts) than the one which has historically affected less advantaged groups. Contributing to this trend will be the growing need to adapt all procedures related to the entire production processes to the automation typical of increasingly AI-managed work environments.

In this sense, control, surveillance and the system of incentives and sanctions (as discussed in Sect. 2 above) will also have to conform to quantification and datafication. The worker's margins of freedom will then likely be reduced as result of the need to conform to strictly quantitative criteria in their actions and in light of the need to be evaluated with tools that prioritize objectivity and efficiency. Ironically, such algorithms are already actively used in the criminal justice system of certain countries (Custers et al., 2022). Indeed, it is hard to see why we could rely on programs that assess the appropriate sentence for an offender, or the possibility of recidivism of the same, after a certain period of imprisonment and not do so for labour disputes.

Another issue relevant to the economic field and to the workers' freedom concerns the possibility of being evaluated and judged by peers and not by AI algorithms (Ernst and Young. 2018; Keystone Consulting, 2017). It is generally agreed that there is a duty of dignity to be accorded to human beings, who should be treated as unique individuals defined by personal traits, and not as a set of data unified by the attribution to a first and last name. This duty of dignity seems to be threatened by the widespread adoption of such technologies.

Another consequence for the workers within the wrapped datafied/infospheric environment in which we increasingly live could be the progressive loss of the freedom to change the rules that govern the environment itself. This is a discretionary activity that does not violate quality standards but allows for changes and improvements in the production process, both technically and in terms of working relationships and conditions. For example, introducing a moment of confrontation between workers can improve both productivity and employee motivation. If, however, the procedures do not allow this, any momentary slowdown in the process will be evaluated negatively, even though it may yield better results in the long run. An efficiency-bound environment that monitors all processes in real time and intervenes to make them homogeneous and smooth cannot tolerate unanticipated deviations and tends to discourage or suppress them.

In this vein, it can be considered another form of freedom: the idea of self-government (Pettit 2011). This latter entails an overall ability to do and not to be governed by alien forces, and the self-mastery ability that sustains the full optionality (or the freedom to do otherwise as specified so far). These two accounts of freedom are logically separated and can vary independently of one another. Now suppose a situation in which we have a high level of optionality, but an environment in which there is a predominance of heteronomy (freedom in the self-governing sense is not respected). In this situation options would be left open to agents, but the agents would not be free by simply having a set of options open, since the algorithms would be in a position to filter the “choice environment” (Danaher 2019).

Thus, in wrapped datafied/infospheric working environments there may be cases with high optionality but with low autonomy. For example, soft control mechanisms over workers' routine, including persuasive pop-ups ads or nudging techniques, have been employed by Uber for steering drivers to have diverse booking options and more flexibility (Scheiber 2017; Webster 2020). Some have noted that this can lead to power asymmetries and structure control over workers.

This is not only the case in strictly structured fields of work such as logistics, but also in fields where AI is only now appearing: such as medical diagnoses, marketing, or the entertainment industry. In all these cases, workers' freedom in decision-making might be reduced in parallel with the possibility to exercise their creativity, as we shall see in the next section.

If in many areas, the human contribution cannot be (yet) dispensed with, one issue related to the progressive depersonalisation of the worker within a datafied environment is that of the loss of the possibility of cultivating and exercising those characteristics in humans that have been shaped

by natural evolution and that AI tends to counteract or suppress (Malinetsky and Smolin 2021). For example, sociality and relationships; the ability to frequent natural and not just artificial environments; other activities including those oriented to a relevant, concrete, and visible purpose.

These aspects are related to physical and mental well-being, which go beyond the immediate gains that the new AI-based economy may bring about in terms of physical security, education, income, or general wealth (even assuming an optimistic scenario, on which many don't necessarily agree). Humans are proactive creatures who deeply fear loneliness, boredom, and feelings of worthlessness. In general, the sense of agency and being held accountable for their actions is something that underlies freedom as a value, as a property that gives meaning to existence from a phenomenological point of view (Farina et al. 2022a).

4 Creativity

Recently, an American artist won the first place in the emerging artist division's "digital arts/digitally-manipulated photography" category at the Colorado State Fair Fine Arts Competition⁵. His winning image, titled "Théâtre D'opéra Spatial," was made with Midjourney⁶- an artificial intelligence system that can produce detailed images when fed written prompts. The affair caused controversy because the (human) jury evaluated the work without considering that it was produced with an AI system; even though the artist openly declared that he had used an AI tool to generate the image upon submitting his work. After the artist got the prize, he was inundated with criticism from numerous colleagues, who deemed it inappropriate to compete with a work made that way. It's like admitting robots to the Olympics, was one of the comments.

There are numerous programs that allow people to create images based on verbal instructions (such as DALL-E 2⁷). Such programs draw on vast image repositories and modify or mix pre-existing figures based on users' inputs. Until now, they were considered curious pastimes, but their entry into competitions and the art market could revolutionise the criteria of creativity, the way it is evaluated, and the role of human beings in contributing to society's creative processes.

Another experiment sparked discussion in early 2022. Two scholars have, with Daniel Dennett's permission and cooperation, "fine-tuned" GPT-3, (the autoregressive language model that uses deep learning to produce human-like

text) on most of Dennett's corpus, with the aim of seeing whether the resulting program could answer philosophical questions similarly to how Dennett himself would. The result was that philosophy experts were unable to clearly discriminate between the answers given by Dennett and answers given by GPT-3⁸.

The two examples we discussed above are just paradigmatic instances of an ongoing revolution focussing on the "creative" possibilities of AI (Miller 2019). The topic of creativity and its definition is one of the most complicated in the field of psychology, but it has to do with the ability to produce something that is new (original and unexpected) and useful (appropriate to the performance of a task) (Sternberg and Lubart, 1999, p. 3). In other words, what is creative is the result of a process that is not necessarily reducible to the mechanics of deterministic reasoning. Usually within creative acts, one cannot identify a precise concatenation of stages but rather perceives holistically the emergence of the result (Koestler, 1964). In contrast, as far as AI creativity is concerned, the operation of the algorithm is potentially "transparent"; that is reducible to a finite number of steps, and its success relates either to the direct liking of a human viewer (as in the case of the art contest mentioned earlier) or in its appropriateness (as in the case of the Dennett-like responses of GPT-3; or other reproductive applications of AI, which thanks to huge databanks and vastly superior computational power to humans can produce a very large number of solutions to a problem, among which to find the most appropriate one).

One may wonder whether we will end up delegating all creative tasks to algorithms, especially in wrapped and data-driven environments, where AI can deploy its engineering capability to the fullest degree. And -if this is the goal-whether human creativity at work will be less and less used. Is this a likely scenario? And what consequences might it entail? Firstly, one may ask whether low-cost, AI-produced creativity is sufficient to meet the needs of consumers (of goods and cultural products) and the resolution of problems that may arise from time to time. Today's computers are composing music that sounds "more Bach than Bach," turning photographs into paintings in the style of Van Gogh's *Starry Night*, and even writing screenplays (Miller 2019). The key point; however, seems to be this: every relevant problem that is more than just a procedural query has to do with humans and their complexity. For example, there is a need to not only save energy and reduce climate-altering emissions, but there is a necessity to do this in tune with the desires and goals of people living in that specific area with a specific culture and specific values.

⁵ <https://arstechnica.com/information-technology/2022/08/ai-wins-state-fair-art-contest-annoys-humans/>, Last Accessed April 2023.

⁶ <https://www.midjourney.com/home/>, Last Accessed April 2023.

⁷ <https://openai.com/dall-e-2/>, Last Accessed April 2023.

⁸ <https://schwitzsplinters.blogspot.com/2022/07/results-computerized-philosopher-can.html>, Last accessed April 2023.

And the same goes for creativity. If we delegate the entire creation and all marketing of -say- a business to a highly efficient algorithm, will so-called creative workers lose their role and over time we will have no more reserves of human creativity? This seems to be related to a certain approach maintaining (perhaps naively) that a “parallel computer” (such as our brain) is capable to produce in ways that are not yet well understood and that exceed the serial capabilities of an analogic computer. However, recent progress on evolutionary computation, especially those grounded on population-based search techniques, seem to suggest the possibility for AI tools (based on parallel processing) to find creative solutions to very practical problems of the real world (Miikkulainen 2021). Evolutionary computation, especially if complemented by deep learning (Schmidhuber 2015; LeCun et al. 2015) can process data both synchronically (in parallel) and diachronically (evolutionarily). It has been observed that population-based search methods based on evolutionary computation can scale better than other machine learning approaches (Miikkulainen 2021, p.163). This suggests that soon we should see many applications of these AI tools to problems directly involving human creativity in numerous fields, such as engineering (Dupuis et al. 2015), healthcare (Miikkulainen et al. 2021), finance (Buckmann et al. 2021), or even in agriculture (Johnson et al. 2019).

There is thus a question of whether the gradual reduction of the creative roles entrusted to humans in highly data-driven environments will lead to an increase in overall system efficiency and increasing consumer satisfaction. Or whether, instead, it may leave uncovered an important part of the innovation that proceeds with the single, unpredictable insights of a few individuals of genius. In addition, the relative untapping of the creativity of workers, who have become executors of the new ideas produced by automated systems, could induce a lowering of the motivation and mood of workers themselves, who will become less and less involved in the production (and decision-making) processes; therefore unable to devise answers even to decisions that for now are still entrusted to humans.

5 Efficacy and Instrumental Reason Issues

The goal of efficiency, as mentioned above, drives the creation of new environments in which AI-based technology may prevail. It is not necessary to refer to Marx’s works to consider how relevant the means of production and the relationships between workers and production processes typical of a given era can be in shaping culture and other types of relationships in society. The logic inherent in the increasingly pervasive application of AI across wrapped and

data-driven (infospheric) environments invites an evaluation of criteria of efficiency, timeliness, and replicability as central to the production process and as particularly valued for what they entail on the wealth and welfare side of consumers and of society as a whole. The so-called instrumental reason; that is, adjusting means to predetermined ends to achieve the best possible outcome, may thus become the benchmark for the entire economic sector (Acemoglu and Restrepo 2020).

In principle, humans are still responsible for decisions concerning the ultimate goals and ultimate choices, but easily find that in the wrapped and datafied microenvironments the whole process revolves around the optimal management of quantitative aspects that can be handled by AI. Speculations about algorithms taking over and altering the purposes for which they were created currently remain science fiction scenarios (Floridi 2022). However, what we may witness in the short term is a culture that may be affected by being increasingly placed in the onlife dimension typical of personal devices, characterized by speed, real time, ever-better performance, and minimization of waiting time or expectations. This has its counterpart in the impatience with slowness and qualitative aspects, with a prevalence of phenomenal aspects over cognitive ones, which distinguish each individual.

Consciousness qua basic feeling of existence, as a background that qualifies all our waking states, seems to be exhibited by at least some living species and, as far as we know, especially by human beings. This is a feature that cannot be replicated or simulated -to date- in artifacts, which however, can be partially exhibited in software as a selective intelligence, sometimes superior, to that of human beings. This is demonstrated by the ability of computers to defeat humans in chess (such as the case with Deep Blue and Kasparov) and even in GO (an abstract strategy board game, where two players play in the attempt to surround more territory than the opponent). These examples show that appreciation for highly developed forms of intelligence also favours the illusion of seeing consciousness where there is none (as in some types of software, e.g., the one in the movie *Her*, with which the protagonist falls in love) and not seeing consciousness where instead it exists (as in non-responsive individuals) (Lavazza and Massimini 2018).

If we pursue forms of intelligent functionalism (cf. López-Rubio 2018), we might end up morally devaluing the criterion of the presence of consciousness in favour of the presence of intelligence-or at least of full consciousness associated with the ability to exercise intelligent functionalism (Lanier 1995). One can, of course, argue in favour of an ethical position of this kind, but it is not easy to do so without completely giving up moral intuition, even in rationally supervised forms. In fact, moral intuition is what seems to

provide us with the basic preconditions of moral reasoning; that is, the fact of sharing at least some of the basic values of the subjects involved (Audi 2015). The latter fact is mainly due to the fundamental quality of living beings: consciousness. And consciousness is something that intelligent artifacts seem to lack, even though they can mimic moral reasoning at a cognitive level.

One consequence of this shift toward quantification, efficiency, speed, and continuous connection is the projection of these machine characteristics to which we have become increasingly accustomed onto our fellow human beings. Tolerance for those who are lower performing, less able to keep up with the pace of the AI systems of which we are gradually becoming a part may be diminishing, starting precisely in workplaces built around automation and possibly extending to the wider society (for instance, in terms of systems revolving around social credit, which may also be based on work performance) (Shew 2020; Nakamura 2019). In those contexts, predictability and reliability are prioritized and measurement ranks first among the system's capabilities. What does not fit within the parameters, what slows down or hinders the flow of the process will tend to be pushed aside, expelled, or not even recruited.

There are several levels at which this selection based on efficiency and instrumental reason can take place. There is a more trivially physical one: those who cannot handle the pace of automation cannot participate in the work process. People affected by different forms of illness or disability, the elderly, and those who fall below minimum performance standards will have difficult access to the labour market and, more importantly, may be seen as less useful to society at large, reversing a trend toward inclusion that has been taking hold recently Stypinska 2022; Farina and Lavazza 2022a, b,c; Farina and Lavazza 2021a).

The same, and perhaps to a greater extent, may happen at the cognitive level, as pointed out earlier. The inability, for various reasons, to keep up and be deeply attuned to the wrapped around and datafied environment could lead to the marginalization of those who manifest such detachment from the new AI-colonized context. This is not an inevitable outcome, but it is a risk that can already be glimpsed in a push for a "digital uniformity" that comes from the now compulsory reliance on electronic devices and indeed even social media with varying forms of indirect control and public exposure.

Heßler and colleagues (2022) noticed that "increased importance of empathy and autonomy leads to a higher degree of algorithm aversion. At the same time, it also leads to a stronger preference for human-like decision support, which could therefore serve as a remedy for an algorithm aversion induced by the need for self-humanization". In recent lab experiments Fuchs (*in press*) found "that the use

of algorithmic (vs. human) management reduces prosocial behavior (e.g., the tendency to help other workers)". In addition, "negative effect (i) occurs because the use of algorithms to manage workers leads to greater objectification of others, (ii) also occurs when algorithms perform tasks together with human managers, and (iii) depends on the type of management task algorithms perform".

Being caught up in the apparent gamification of an increasing number of tasks and functions through digital technology may lead to an overvaluing of instrumental reason at the expense of a search for ends and values to which one can give motivated and thoughtful personal adherence. Muldoon and Raekstad (2022) proposed the concept of "algorithmic domination", where an individual "is subjected to a dominating power, the operations of which are (either in part or in whole) determined directly by an algorithm". Also gamification permits employers "to intervene at a more minute level in ways that are not feasible if required to be undertaken by a human supervisor".

In this scenario, the business sector seems to be destined to be increasingly pervaded by AI. Producing quantifiable, guaranteed, and predictable results is one of the main goals of deeply wrapped datafied environments, a goal that tends to leave no room for uncontrollable and uncontrolled personal paths. Such a scenario, we maintain, requires careful ethical evaluation and constant scrutiny to avoid that a single efficientistic view (incapable of an inclusive look at every human being) may prevail.

6 Conclusion

As AI becomes ubiquitous in society, possibly leading to the formation of increasingly intelligent bio-technological unions, there will likely be a coexistence of a plethora of micro-environments wrapped and tailored around robots and humans. The key element of this pervasive process will be the capacity to integrate biological realms into an info-sphere suitable for the implementation of AI technologies. This process will likely require extensive datafication.

This trend can help to meet an increasing number of needs of a growing share of the population by improving the efficiency of production processes and introducing into them elements of quantification, predictability, reproducibility, and minimization of error and imperfection. All this, however, can also trigger unintended and suboptimal consequences. In this paper we have considered four such consequences that seems to be crucial for decision-making processes in future human societies dominated by AI technologies.

The datafication required to realize the quantification and application of AI resources implies increasing control of the

individual involved in the production process and quite possibly over her life. This loss of privacy is typical of new datafied and infospheric environments, where it is not necessarily pursued with the explicit purpose of monitoring the individual (*surveillance*) but rather of actively predicting her behaviour (*dataveillance*), by having the individual herself interacting effortlessly with a wide range of integrated technological tools across multiple domains and dimensions.

This need for control may also result in a loss of freedom understood in the classical sense as the possibility of deciding between alternative courses of action. Indeed, the worker must manifest maximally predictable behaviour for her contribution to be as effective and integrated as possible. Freedom, in this context, may become structurally endangered as an end-product, especially in production processes, which are increasingly oriented toward maximal certainty (which is the opposite of freedom).

Another consequence of this production arrangement is the delegation of creativity to algorithms, which are often presented as higher performing, hence preferable to humans because -unlike humans- they are not subject to quantitative and qualitative fluctuations. The risk here is a loss of the reserve in terms of qualitative resource on the part of workers, which -in the long term- could leave some creative areas uncovered, especially those where machines are not (yet) at the level of productive intelligence of humans.

Finally, a more general cultural tendency to favour efficiency and instrumental reason might assert itself because of the structural constraints that environments wrapped around AI tend to produce. A less inclusive and tolerant society could be the result of our onlife characterized by immediacy and absence of expectations, a world -in brief- where common-sense will leave space to pure objectivity and absolute neutrality based on algorithmic efficiency.

In this vein, datafication points toward an automation of decision-making that makes it primarily efficiency-driven toward predetermined goals. One strategy to rebalance this trend could be to create areas of decision-making that are removed from extreme datafication to allow a process of decision making driven by the choices of individuals without the close guidance of AI. For we know that a strong sense of agency is inherent in human beings, consisting of (presumed) conscious control over their choices and courses of action. The deprivation of this sense of agency usually leads individuals to a reduction of their own well-being (Creed and Klisch 2005).

If, therefore, the efficiency of economic organization is not to become the first and only goal of the social system, with the consequences just highlighted, it is necessary to prevent a form of automated decision-making (based on datafication) from becoming the only method for choices in working environments and in societies in general. Humans

are quintessentially social beings, who are bound to have contacts with their peers to find satisfaction, often in free and unstructured interactions. The cancellation of these interactions can trigger a reduction in their well-being far greater than the support they could get from the intelligent tools located in increasingly datafied environs.

So, as mentioned above, potential risks exist that need to be addressed pre-emptively as they seem to be inherent in structural trends (and aspects of decision-making processes) based on the widespread diffusion of AI in society. It is the task of philosophy and ethics to help analyse these risks, highlight their contours, and propose solutions so that artificial intelligence may become a valuable complement to human activities, favouring (rather than hampering) social harmony and moral good.

Author contributions The authors contributed equally to the writing of this paper.

Funding N/a.

Data Availability N/a.

Declarations

Disclosure of Potential Conflicts of Interest None.

Research involving Human Participants and/or Animals: n/a.

Informed Consent N/a.

Competing Interests N/a.

References

- Acemoglu D, Restrepo P (2020) The wrong kind of AI? Artificial intelligence and the future of labour demand. *Camb J Reg Econ Soc* 13(1):25–35
- Ajana B (2017) Digital health and the biopolitics of the quantified self. *Digit Health* 3:1–18
- Ali W, Tian W, Din SU, Iradukunda D, Khan AA (2021) Classical and modern face recognition approaches: a complete review. *Multimedia Tools and Applications* 80(3):4825–4880
- Alsaadi IM (2021) Study on most popular behavioral biometrics, advantages, disadvantages and recent applications: a review. *Int J Sci Technol Res* 10:15–21
- Anumanchipalli GK, Chartier J, Chang EF (2019) Speech synthesis from neural decoding of spoken sentences. *Nature* 568(7753):493–498
- Audi R (2015) Intuition and its place in ethics. *J Am Philosophical Association* 1(1):57–77
- Bakare A, Bygayenko E, Cheverda A, Farina M, Pedrycz W, Plaskin Y, Succi G (2022) Automatically Prioritizing and Assigning Tasks from Code Repositories in Puzzle Driven Development. *Proceedings of the 19th International Conference on Mining Software Repositories (MSR)*, doi: <https://ieeexplore.ieee.org/document/9796201>
- Ball K, Haggerty K, Lyon D (2012) Routledge handbook of surveillance studies. Routledge, London, UK

- Bauman Z, Lyon D (2013) Liquid Surveillance: a conversation. John Wiley & Sons, Hoboken, NJ
- Beer D (2009) Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media Soc* 11:985–1002
- Benziane S, Benyettou A (2011) An introduction to biometrics. *Int J Comput Sci Inform Secur* 9(4):40–47
- Birdsell M (2014) Google and ITE: the road ahead for self-driving cars. *Inst Transp Eng ITE J* 84(5):36
- Bonnefon JF (2021) The car that knew too much: can a machine be moral? MIT Press, Cambridge, MA
- Borenstein J, Herkert JR, Miller KW (2019) Self-driving cars and engineering ethics: the need for a system level analysis. *Sci Eng Ethics* 25(2):383–398
- Boutros F, Damer N, Kirchbuchner F, Kuijper A (2022) Elasticface: Elastic margin loss for deep face recognition. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1578–1587)
- Buckmann M, Haldane A, Hüser AC (2021) Comparing minds and machines: implications for financial stability. *Oxf Rev Econ Policy* 37(3):479–508
- Bugayenko Y, Bakare A, Cheverda A, Farina M, Kruglov A, Plaskin Y, Pedrycz W, Succi G (2023) Prioritizing tasks in software development: a systematic literature review. *PLoS ONE (Comput Inf Sci)* 18(4):e0283838. <https://doi.org/10.1371/journal.pone.0283838>
- Chamorro-Premuzic T, Akhtar R, Winsborough D, Sherman RA (2017) The datafication of talent: how technology is advancing the science of human potential at work. *Curr Opin Behav Sci* 18:13–16
- Christophersen M, Mørck P, Langhoff TO, Bjørn P (2015) Unforeseen challenges: adopting wearable health data tracking devices to reduce health insurance costs in organizations. In: Antona, M., Stephanidis, C. (Eds.). *International Conference on Universal Access in Human-Computer Interaction*, (pp. 288–299). Berlin, Germany: Springer
- Creed PA, Klisch J (2005) Future outlook and financial strain: testing the personal agency and latent deprivation models of unemployment and well-being. *J Occup Health Psychol* 10(3):51–260
- Cukier K, Mayer-Schoenberger V (2013) The rise of big data: how it's changing the way we think about the world. *Foreign Aff* 92:28
- Custers B (2022) AI in Criminal Law: an overview of AI applications in substantive and procedural Criminal Law. *Law and Artificial Intelligence*, 205–223
- Da Bormida M (2021) The Big Data World: benefits, threats and ethical Challenges. Ethical issues in Covert, Security and Surveillance Research. Emerald Publishing Limited
- Danaher J (2019) The Ethics of Algorithmic Outsourcing in Everyday Life. In: Young K, Lodge M (eds) Algorithmic regulation, (pp 98–118). Oxford, UK: Oxford University Press
- Delfanti A (2021) Machinic dispossession and augmented despotism: digital work in an Amazon warehouse. *New Media & Society* 23(1):39–55
- Dupuis JF, Fan Z, Goodman E (2015) Evolutionary design of discrete controllers for hybrid mechatronic systems. *Int J Syst Sci* 46(2):303–316
- Ellerbek A (2011) Playful biometrics: controversial technology through the lens of play. *Sociol Q* 52:528–547
- Engstrom E, Jebari K (2022) AI4People or People4AI? On human adaptation to AI at work. *AI & Society*, 1–2, <https://link.springer.com/article/10.1007/s00146-022-01464-5>
- Ernst and Young (2018) The New Age: artificial intelligence for Human Resource Opportunities and Functions. Ernst and Young LLP, London
- Farina M, Lavazza A (2021a) Knowledge prior to belief: is extended better than enacted? *Behav Brain Sci* 44. <https://doi.org/10.1017/S0140525X2000076X>
- Farina M, Lavazza A (2021b) The meaning of Freedom after Covid-19. *Hist Philos Life Sci* 43(1):1–5
- Farina M, Lavazza A (2022a) Mind Embedded or Extended: Transhumanist and Posthumanist Reflections in Support of the Extended Mind Thesis. *Synthese*, 200, 507, doi: <https://doi.org/10.1007/s11229-022-03963-w>
- Farina M, Lavazza A (2022b) Why there are still moral reasons to prefer extended over embedded: a (short) reply to Cassinadri. *Philos Technol* 35:67. <https://doi.org/10.1007/s13347-022-00566-8>
- Farina M, Lavazza A (2022c) Incorporation, transparency, and cognitive extension. Why the distinction between embedded or extended might be more important to ethics than to metaphysics. *Philos Technol* 35:10. <https://doi.org/10.1007/s13347-022-00508-4>
- Farina M, Gorb A, Kruglov A, Succi G (2022a) Technologies for GQM-based metrics recommender systems: a systematic literature review. *IEEE Access* 10:23098–23111. <https://doi.org/10.1109/ACCESS.2022.3152397>
- Farina M, Karimov A, Zhdanov P, Lavazza A (2022b) AI and society: a Virtue Ethics Approach. *AI Soc*. <https://doi.org/10.1007/s00146-022-01545-5>
- Floridi L (1999) Philosophy and Computing: an introduction. Routledge, New York
- Floridi L (2011a) Children of the fourth revolution. *Philos Technol* 24(3):227–232
- Floridi L (2011b) Enveloping the World for AI. *The Philosophers' Magazine* 54:20–21
- Floridi L (2013) The Ethics of Information. Oxford University Press, Oxford
- Floridi L (2014) The fourth revolution: how the Infosphere is reshaping human reality. Oxford University Press, Oxford
- Floridi, L. (2022). Etica dell'Intelligenza Artificiale. Raffaello Cortina Editore, Milan, Italy
- Fotopoulou A, O'Riordan K (2016) Training to self-care: fitness tracking, biopedagogy and the healthy consumer. *Health Sociol Rev* 26:54–68
- Foucault M (1991) Governmentality, transl. R. Braidotti. In: Burchell G, Gordon C, Miller P (eds) *The Foucault Effect: studies in Governmentality*. University of Chicago Press, Chicago, IL, pp. 87–104
- Foucault M (1997) *Society Must Be Defended: Lectures at the College de France, 1975–1976*. New York, NYC: St. Martin's Press
- Fuchs C (in press). Management by Algorithms reduces Prosocial Behavior
- Gitelman L (2013) Raw Data is an Oxymoron. MIT Press, Cambridge, MA
- Goretzko D, Israel LSF (2021) Pitfalls of machine learning-based Personnel Selection. *J Personnel Psychol* 21:37–47
- Gray D, Henderson SE (2017) *The Cambridge handbook of surveillance law*. Cambridge University Press, Cambridge, UK
- Gstrein OJ, Beaulieu A (2022) How to protect privacy in a datafied society? A presentation of multiple legal and conceptual approaches. *Philos Technol* 35(1):1–38
- Heaven WD (2022) This startup's AI is smart enough to drive different types of vehicles. *MIT Technology Review*, <https://www.technologyreview.com/2022/09/26/1060023/wayve-ai-driverless-vehicles-cars-vans-scale-up/#:~:text=Wayve%2C%20a%20driverless%2Dcar%20startup,learned%20to%20drive%20multiple%20vehicles.>
- Heßler PO, Pfeiffer J, Hafenbrädl S (2022) When Self-Humanization leads to Algorithm Aversion. *Bus Inform Syst Eng* 64(3):275–292
- Hoeyer K, Wadmann S (2020) “Meaningless work”: how the datafication of health reconfigures knowledge about work and erodes professional judgement. *Econ Soc* 49(3):433–454
- Hu M (2017) Biometric surveillance and big data governance. In: Gray D, Henderson S, Eds (eds) *The Cambridge Handbook of Surveillance Law*. Cambridge University Press, Cambridge, UK, pp. 121–149

- Jain AK (2007) Biometric recognition: overview and recent advances. Iberoamerican Congress on Pattern Recognition. Springer, Berlin, Heidelberg, pp. 13–19
- Jarke J, Breiter A (2019) Editorial: the datafication of education. *Learn Media Technol* 44(1). <https://doi.org/10.1080/17439884.2019.1573833>
- Johnson AJ, Meyerson E, de la Parra J, Savas TL, Miikkulainen R, Harper CB (2019) Flavor-cyber-agriculture: optimization of plant metabolites in an open-source control environment through surrogate modeling. *PLoS One*, 14(4), e0213918
- Kennedy H, Poell T, van Dijck J (2015) Introduction: data and agency. *Big Data Soc* 2. <https://doi.org/10.1177/2053951715621569>
- Keystone C (2017) Big-data and its role on the efficacy of human resource management. Keystone Consulting, Investigative Report, Shanghai
- Lanier J (1995) Agents of alienation. *J Conscious Stud* 2(1):76–81
- Lavazza A, Farina M (2021) Experts, naturalism, and democracy. *J Theory Soc Behav* 52(2):279–297. <https://doi.org/10.1111/jtsb.12321>
- Lavazza A, Inglese S (2015) Operationalizing and measuring (a kind of) free will (and responsibility). Towards a new framework for psychology, ethics, and law. *Rivista internazionale di Filosofia e Psicologia* 6(1):37–55
- Lavazza A, Massimini M (2018) Cerebral organoids and consciousness: how far are we willing to go? *J Med Ethics* 44(9):613–614
- LeCun Y, Bengio Y, Hinton G (2015) Deep Learn Nat 521(7553):436–444
- López-Rubio E (2018) Computational functionalism for the deep learning era. *Mind Mach* 28(4):667–688
- Lupton D (2016) The quantified self: a sociology of Self-Tracking. Polity Press, Cambridge, UK
- Lyon D (2014) Surveillance, Snowden, and big data: capacities, consequences, critique. *Big data & society* 1(2):2053951714541861
- Mackenzie A (2005) The performativity of code: software and cultures of circulation. *Theory Cult Soc* 22:71–92
- Malinetsky G, Smolin V (2021) The artificial intelligence influence on real sociality. *Procedia Comput Sci* 186:344–351
- Mayer-Schönberger V, Cukier K (2013a) Big data: a revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt
- Mayer-Schönberger V, Cukier K (2013b) *Big data: the essential guide to work, life and learning in the age of insight*. Hachette UK
- McCarthy J, Minsky ML, Rochester N, Shannon CE (2006) A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI magazine* 27(4):12–12
- McQuillan D (2016) Algorithmic paranoia and the convivial alternative. *Big Data Soc* 3. <https://doi.org/10.1177/2053951716671340>
- Miikkulainen R (2021) Creative AI through evolutionary computation: principles and examples. *SN Comput Sci* 2(3):1–7
- Miikkulainen R, Francon O, Meyerson E, Qiu X, Sargent D, Canzani E, Hodjat B (2021) From prediction to prescription: evolutionary optimization of nonpharmaceutical interventions in the COVID-19 pandemic. *IEEE Trans Evol Comput* 25(2):386–401
- Miller AI (2019) The artist in the machine: the world of AI-powered creativity. MIT Press, Cambridge, MA
- Muldoon J, Raekstad P (2022) Algorithmic domination in the gig economy. *Eur J Political Theory*. <https://doi.org/10.1177/14748851221082078>
- Nakamura K (2019) My algorithms have determined you're not human: AI-ML, reverse turing-tests, and the disability experience. *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, <https://doi.org/10.1145/3308561.3353812>
- Nixon MS, Correia PL, Nasrollahi K, Moeslund TB, Hadid A, Tistarelli M (2015) On soft biometrics. *Pattern Recognit Lett* 68:218–230
- Pettit P (2011) The instability of freedom as noninterference: the case of Isaiah Berlin. *Ethics* 121(4):693–716
- Pietrini P, Lavazza A, Farina M (2022) Covid-19 and biomedical experts: when epistemic authority is (probably) not enough. *J Bioethical Inq* 19:135–142. <https://doi.org/10.1007/s11673-021-10157-5>
- Pizza M, Pecetta S, Rappuoli R (2021) Vaccines 2020: the era of the digital vaccine is here. *Sci Transl Med* 13(624):eabm3249
- Pybus J, Coté M (2021) Did you give permission? Datafication in the mobile ecosystem. *Information, Communication & Society*, 1–19
- Ramachandra R, Busch C (2017) Presentation attack detection methods for face recognition systems: a comprehensive survey. *ACM Comput Surv (CSUR)* 50(1):1–37
- Ruckenstein M, Schüll ND (2017) The datafication of health. *Annu Rev Anthropol* 46:261–278
- Ruppert E (2011) Population objects: interpassive subjects. *Sociology* 45:218–233
- Ryu R, Yeom S, Kim SH, Herbert D (2021) Continuous multimodal biometric authentication schemes: a systematic review. *IEEE Access* 9:34541–34557
- Scheiber N (2017) How Uber Uses Psychological Tricks to Push Its Drivers' Buttons. *New York Times*, April 02. <https://www.nytimes.com/interactive/2017/04/02/technology/uber-drivers-psychological-tricks.html>
- Schmidhuber J (2015) Deep learning in neural networks: an overview. *Neural Netw* 61:85–117
- Shew A (2020) Ableism, technobableism, and future AI. *IEEE Technol Soc Mag* 39(1):40–85
- Skenderija S, "Miroslav Petricek and the Quest for a New Ontology of Information" (2008) *Cornell Law Faculty Working Papers*. 34. <https://scholarship.law.cornell.edu/clspapers/34>
- Smith M, Miller S (2022) The ethical application of biometric facial recognition technology. *AI Soc* 37(1):167–175
- Stypinska J (2022) AI ageism: a critical roadmap for studying age discrimination and exclusion in digitalized societies. *AI Soc* 1–13. <https://doi.org/10.1007/s00146-022-01553-5>
- Till C (2017) Commercialising bodies: the new corporate health ethic of philanthrocapitalism. In: Lynch R, Farrington C (eds) Quantified lives and vital data. Palgrave-Macmillan, Basingstoke, UK, pp. 229–249
- Tippins NT, Oswald FL, McPhail SM (2021) Scientific, legal, and ethical concerns about AI-based personnel selection tools: a call to action. *Personnel Assess Decisions* 7(2):1. <https://scholarworks.bgsu.edu/pad/vol7/iss2/1/>
- Vacca JR (2007) Biometric technologies and verification systems. Elsevier, Butterworth-Heinemann, UK
- Van Dijck J (2014) Datafication, dataism and dataveillance: big data between scientific paradigm and ideology. *Surveill Soc* 12:197–208
- Webster E (2020) The Uberisation of work: the challenge of regulating platform capitalism. A commentary. *Int Rev Appl Econ* 34(4):512–521
- Williamson B, Bayne S, Shay S (2020) The datafication of teaching in Higher Education: critical issues and perspectives. *Teach High Educ* 25(4):351–365
- Woods SA, Ahmed S, Nikolaou I, Costa AC, Anderson NR (2020) Personnel selection in the digital age: a review of validity and applicant reactions, and future research challenges. *Eur J Work Organizational Psychol* 29(1):64–77
- Yin S, Kaynak O (2015) Big data for modern industry: challenges and trends [point of view]. *Proceedings of the IEEE*, 103(2), 143–146
- Zuboff S (2019) The age of surveillance capitalism: the fight for a human future at the New Frontier of Power. Profile Books, London, UK

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted

manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.