Simona Tiribelli[*]

Artificial Intelligence and Moral Progress: A Missing Debate

Abstract
This article addresses an underexplored topic in the field of ethics of artificial intelligence (AI), that is, the impact of AI technology on moral progress at the collective level, and shows that such an inquiry is not only ethically relevant but necessary for the design of AI systems able to protect our moral gains and sustain the possibility of further societal moral progress in our contemporary information societies increasingly shaped if not yet governed by AI technology.

Keywords
Ethics of Artificial Intelligence (AI); Societal Moral Progress; Collective Moral Reasoning; Epistemic Agency; Ethics by Design

1. Moral progress and AI ethics

Over the last decade, a great deal of philosophical work has been focused on identifying, assessing, and addressing the ethical risks and the controversial societal implications of the fast-paced design, deployment, and use of AI technology. The main approach underlying such work, which is structured today in the specific applied ethics field known as AI ethics, is a *proactive one*: it is grounded on the idea that we should develop specific ethical knowledge or apply ethical theories as developed in moral philosophy and applied ethics to prevent AI ethical risks, instead of just mitigating by chasing them afterward, through an approach of *ethics by design*. Prominent research in AI ethics does not only focus on how to design AI systems that prevent further social discrimination and new human rights infringement phenomena. They focus also on how to

[*] Università degli Studi di Macerata.

Filosofia morale/Moral Philosophy, vol. 2, 2024 • ISBN: 9791222318318 • DOI: 10.7413/fmmp0063

proactively ethically design such systems to address socially rooted problems, such as unfair inequalities, and create fairer and *morally better* societies (O'Neil 2016; Benjamin 2018; Giovanola & Tiribelli 2022). Indeed, the goal of a prominent deal of scholarship in AI ethics does not lie in the design of neutral or objective AI technology, which would preserve the *status quo*. This approach would be indeed sterile and inadequate in non-ideal societies as the real ones to prevent morally undesirable outcomes; that is, for example, to avoid or mitigate that AI can reflect, perpetuate, or exacerbate sources of harm such as systemic ethnic bias, structural discrimination, and unfair social inequalities that are historically rooted and culturally embedded into our societies. Rather, it focuses on embedding and translating specific ethical values into AI design, often as *ethics by design* criteria, to develop AI systems that can fix or mitigate such sources of harm and enable only those specific technical progress and advancements that both respect and promote the moral values and social goals we care collectively. That is, only that technical progress that can preserve our moral conquests (e.g., widespread acknowledgment of fundamental rights) and promote further societal and moral progress (e.g., fairer societies through further inclusion, made possible by AI systems, or through the detection and removal of historically-rooted unfair cultural biases, thanks to AI big data processing and analytics techniques)[1].

In sum, it sounds we could rightly claim that a great corpus of AI ethics scholarship assumes – more or less implicitly – moral progress in its background or tends to it as its purpose. Nevertheless, the specific topic of moral progress is surprisingly poorly explored explicitly in this domain. Thus far, a systematic ethical inquiry based on insights from theories on moral progress to investigate whether AI systems promote or hinder the necessary preconditions for moral progress at the collective level has neither been carried out in the AI ethics debate nor in the moral progress one[2].

This article addresses such a gap and shows that such an inquiry is not only ethically relevant but necessary for the design of AI systems in a way to protect our moral gains and avoid phenomena of moral regress in our contemporary mature information societies that are increasingly and deeply shaped – if not yet governed – by algorithmic technology.

To this aim, the paper first clarifies the philosophical account of moral progress we adopt in our ethical inquiry, which is one of today's bench-

[1] See Tiribelli (2023) for a proposal on the use of AI capabilities to detect and mitigate sources and drivers of systemic unfair social inequalities into our societies.

[2] Indeed, just a few (precious) attempts have been made with regard to AI systems and moral progress that however focus on moral change at the individual level; see the analyses of Savulescu & Giubilini (2018) and Savulescu & Maslen (2015).

mark theories on moral progress, Buchanan and Powell's biocultural account of moral progress, and specifically the necessary preconditions for moral progress such account points out. Then, the paper shows how the AI systems hyperconnected into our information societies are impacting such preconditions, zooming in on their impact on the widespread exercise of open-ended moral reasoning, sketching out a few implications such impact raises. Finally, the paper concludes by calling for specific ethics by design criteria to shape the development of such AI systems so as to protect and promote the necessary preconditions securing at least at a minimum threshold the possibility for further moral progress and the basis for avoiding moral regress.

2. On moral progress

The concept of moral progress is a complex one and it is widely debated within a plurality of domains, from moral philosophy, through evolutionary psychology, to biology and cognitive (neuro)sciences. It is not within the scope of the present article to debate this concept or provide a related systematic discussion as it would lead us astray. For the purpose of the paper, however, we aim here to briefly clarify the (i) conception of moral progress we endorse in our ethical analysis considering the general debate on moral progress, (ii) the diverse types of moral progress in this sense we can contemplate, the (iii) specific account of moral progress we encompass for our inquiry, and (iv) the particular *sine qua non* preconditions for its possibility such account contends.

As it has been pointed out by Sauer et al. (2021), it is possible to distinguish at least *two* major approaches within the contemporary debate that mainly emerge on moral progress. One approach is of those who embrace a *broad conception* of moral progress, according to which moral progress can be understood as "any kind of morally desirable change" (Sauer et al. 2021, p. 2; Eriksen, 2020, pp. 5-11; Kitcher, 2017, p. 64; Sauer, 2019) from individual changes in beliefs and behaviors (e.g., a mother who learn how to be more patient with their children) to progress on a global scale (a global decline in homicide rates) to improvements in moral theory. Another approach is, instead, of those who endorse a *narrow conception* of moral progress: these scholars content that not all morally desirable changes should be considered as moral progress, but only those that are determined, for example, by the exercise or improvements of human moral capacities (Buchanan & Powell 2018, p. 51), such as moral reasoning, moral motivation, or the ability to follow moral norms (think of those moral gains in human welfare that depend on the large-scale spread of ideas about fundamental rights within some societies). In

this article, we endorse a *narrow* conception of moral progress, according to which we can consider instances of moral progress only those morally loaded phenomena that substantially depend on – or largely involve – the exercise of human *moral capacities* by a large number of people within a society. In this sense, as anticipated in the introduction of this paper, we are interested here in moral progress at the collective level, that is, moral changes that concern a widespread number of people (Buchanan & Powell 2018; Singer 1981); a number that is sufficient to elicit socio-political revolutions as those that have enabled in human history those moral conquests we refer today also as benchmark or historical examples of moral progress (e.g., the abolition of slavery)[3]. We specifically embrace this second view insofar as it entails the *active involvement* of the human *moral dimension* at the collective level and here we are specifically interested in understanding how AI systems can impact this human *moral dimension* and related implications for moral progress. Put it differently, we are not interested on investigating if the surge in application of AI systems may accidentally reduce some morally detrimental phenomena for reasons that have nothing to do with the exercise of practical agency of individuals and groups. Rather, we are specifically interested in investigating whether AI systems currently in use are designed and hence "operate" in such a way to promote or instead undermine the exercise of human moral capacities, such as moral reasoning, moral motivation, and the normative agency and, by doing so, affect moral progress.

In this (narrow) sense, classic examples of moral progress are historically traced, with broad agreement in the debate, to the abolition of slavery, the establishment, at least in some parts of the world, of universal suffrage, and the advancement of LGBTQ+ and animal rights. It is possible to distinguish more specific types of moral progress in its narrow conception, that is, as dependent on the exercise of human moral capacities. These types may consist of adequate *de-moralizations* that concern wrongly moralized practices in the past (e.g., sex outside marriage), or instead in *correct moralizations* of behaviors erroneously thought to be morally neutral (e.g., gender-based wage disparities); or in cases of *improvements in moral motivation* (when this comes to play a determining role on people's behavior, leading to greater levels of compliance with moral norms) as well as in the *improvement of existing moral concepts* (for instance, the concept of responsibility thanks to that of intentionality) and in the *development of new moral concepts* (e.g., the concept of

[3] We do not discuss here the nature of moral progress (if it is global, that is, only happens at the level of institutions, etc., or applies to the level of individuals) or mean it in global terms (i.e., for all humans and societies worldwide); we always refer to it here in a *localized* way (i.e., as it happens in certain socio-political contexts and societies).

sexual harassment), which enable people to recognize, understand, and communicate in an intelligible way to others a certain kind of injustice they are subject to (Fricker 2007)[4].

Most of these types of moral progress are discussed against the backdrop of one of the most endorsed theories of moral progress in the contemporary debates, Buchanan and Powell's biocultural theory of moral progress, which is situated within the classical view of moral progress as inclusion, or gradual expansion of the circle of moral consideration (Buchanan & Powell 2018; Singer 1981). This theory is of particular interest here and argues for the possibility of moral progress and regression as dependent on the dynamic interaction between particular social environmental conditions (and the *stimuli* they generate) and the flexible or ductile (moral) cognitive capacities of individuals, which are adaptively plastic. In short: based on the stimuli that different social environments can provide, people can develop inclusive or, instead, divisive (or tribalistic) moral responses (especially morally regressive outgroup hostility) and so promote further moral progress or instead trigger phenomena of moral regress, as moral progress and conquests are not irreversible.

In particular, according to Buchanan and Powell's account, the presence of favorable socio-epistemic and institutional conditions would enable the exercise of a fundamental *necessary condition* in order to sustain moral progress[5], in terms of large-scale socio-political changes, as those mentioned above: the *widespread exercise of critical, open-ended moral reasoning* (Buchanan & Powell 2018; Buchanan 2021, p. 44), whereby it is meant the ability of cognitively normal human beings have to make the particular moral rules and concepts they are following objects of critical scrutiny, which can sometimes lead to their modification or abandonment (Buchanan 2021, pp. 142-143). In summary, it is the "capacity to examine our moral concepts and the concept of morality itself" (Buchanan 2021, pp. 144-145)[6]. By contending how the plasticity of human cognitive capacities under favorable social (ecological) conditions can foster moral reasoning and prompt inclusive moral responses at large-scale, hence sustaining moral progress, this account acknowledges both the contributions of empirical sciences and evolutionary biology, as well as those of neuroscience (findings on cognitive bias, limits, and errors), while leav-

[4] We refer to the work of Sauer et al. (2021) for a systematic discussion on different types of moral progress.

[5] We take these conditions as what is required at a minimum threshold for safeguarding the chance of further moral progress; the question of whether they are also sufficient is not discussed here.

[6] For key insights on reflexivity, consider theorists such as (classically) Horkheimer (see Traditional and Critical Theory, 1937) or, more recently, Jaeggi (Critique of Forms of Life, 2018).

ing room for the critical role of *human moral normativity* (Buchanan & Powell 2018, pp. 145-155).

Furthermore, this theory has the merit of being among the few to illuminate certain *sine qua non* conditions for moral progress at the collective level, that is, in terms of large-scale socio-political changes (i.e., involving moral improvements of a sufficiently large number of people). In short, these preconditions for moral progress are identified, first, in certain favorable *socio-epistemic conditions*; we sum up briefly a few of them as follows. There should be (i) the presence of advanced technologies to disseminate ideas and connections between ideas that enable reasoning about how people ought to live; (ii) there must be considerable freedom of expression and association so that discussion of ideas about how to live among people with diverse viewpoints can be relatively open-ended and the control over information and communication technologies (ICT) must be dispersed, so as to avoid constraints on freedom of information and expression and the curtail the exercise of critical moral reasoning; (iii) a significant number of people must be exposed to the fact that other societies have different ways of doing things (this allows for the possibility of investigating whether the moral rules they follow are optimal, the extension of sympathy beyond their group, and the identification of common interests); (iv) a developed culture of justification, of providing genuine reasons, not based, therefore, on appeals to authority or tradition. As such account points out (see Buchanan 2021, p. 149), such necessary socio-epistemic conditions depend on the character of the institutions in a society too and the sort of institutional order in which large-scale moral progress is likely to come about through peaceful means will be in broadest terms a liberal and at least minimally democratic order. Taken together, these pre-conditions are claimed to be enabling a third (moral – strictly speaking) condition, which is contended as key to the achievement of moral progress: the widespread exercise of open-ended, critical moral reasoning. Put differently: they constitute the social-epistemic contexts that are deemed necessary for the large-scale exercise of our moral capacities, and therefore, for the possible occurrence of the various types of moral progress in a narrow sense previously mentioned. For the sake of clarity, we do not proceed further on moral progress: we have indeed enough elements to carry out our ethical analysis.

In this section, we have shed light on the specific account of moral progress and the *sine qua non* conditions of moral progress we refer in order to carry out our ethical inquiry, with a specific focus on the widespread exercise of moral reasoning. In the next section, drawing on these insights, we analyze how AI systems and particularly algorithmic ICT can impact such preconditions and the implications for moral progress such impact raises.

3. Moral progress and AI systems

The impact of AI technology on the social and epistemic conditions of our agency in mature information societies has been extensively documented within the scholarship on digital ethics and ethics of AI. Early scholarship in digital ethics had already outlined how digital ICTs were reshaping the environments where we live, renewing them as novel hyperconnected, phygital, onlife ecosystems we are fully and constantly immersed in, as well as the social practices through which we develop, process, and communicate knowledge about the reality, the others, and ourselves (Floridi 2014). Later, further work has refined and expanded these arguments with special regard to the pervasive use of machine learning (ML) and deep learning (DL) algorithms that nowadays govern almost every digital ICT, from search engines to social networking services (large language model-based platforms included) and Internet of things and smart devices broadly.

The idea is catching on in this debate claims the rise of an "algorithmic governance" in today's information societies (Yeung 2018; Tiribelli 2022), according to which ML and DL algorithms are restructuring and predefining the information-based social and epistemic environments in which we prepare and make our choices as particular individuals and as belonging to social groups. They are no longer merely "information gatekeepers" but are playing the role of new kind of choice architects (Tiribelli 2022). This is possible because, due to the ubiquitous presence of algorithm-based digital ICT, everything about reality and ourselves, from our identity characteristics, features, and onlife movements to our relationships, social affiliations, most intimate attachments, sexual orientations, attitudes, preferences, beliefs, vulnerabilities, values, and shared commitments, cannot be only described in informational terms, but can be datafied and hence practically translatable (as captured, processed, incorporated) into data and information. Everything – what we produce and what feeds us – is data and information, and the latter is processed and managed today by AI algorithms such as ML and DL.

This means that the algorithms underlying our ICTs not only decide today what information to show us, and in what order – through filtering and classification techniques – but also determine how (based on which parameters and specific objectives) our onlife *informational socio-epistemic environments* are defined (i.e., the environments where we are immersed and informationally exposed to both knowledge and relationships). In other words, these systems reshape our onlife environments in terms of informational and socio-relational exposure and availability. By doing so, they influence the socio-epistemic and moral conditions of our agency in mature information societies, by choosing what and who – both

quantitatively and qualitatively – will inform and shape the development of our knowledge (i.e., what ideas, thoughts, beliefs, relationships, etc. we will meet and will challenge our thoughts and what won't as instead non-available), including moral knowledge (i.e., what kind of moral reasons, values, and everything that can become a motive for making certain choices and taking specific actions over others will be available and what it won't).

The design of such systems, however, is oriented by parameters and objectives pre-established by third parties (i.e., technology designers and deployers), and as currently conceived, may give rise to phenomena that negatively affect at least a few of the above-mentioned socio-epistemic and moral conditions necessary for moral progress and consequently the possibility for moral progress itself.

Indeed, the ML and DL algorithms at stake here are probabilistic, meaning that they learn how to achieve predetermined goals – which today mostly (sometimes still exclusively) overlap with profit maximization through efficiency criteria – by scaling huge amounts of data to discover knowledge (i.e., patterns, correlations, etc.) that is efficiently used to categorize the users into specific groups of individuals profiled with similar identity and decision-making characteristics including similarities in cognitive distortions and biases (O'Neil 2016). These groups and hence the people silently placed in them are then presented with filtered informational (socio-relational and epistemic) contexts that are personalized for them (e.g., information and relationships aligned to the group profiled major political orientation), but specifically oriented to the efficient achievement of ML and DL goals preset by tech providers, such as maximizing clicks on third-party sponsored content (Zuboff 2018).

In doing so, however, as it has been widely documented over the last decades (Pariser 2011; Sunstein 2008, 2017; Nguyen 2020), these systems elicit some specific socio-epistemic controversial phenomena, such as the formation of filter bubbles, epistemic bubbles, or informational silos (echo-chambers), situating individuals within groups of like-minded, i.e., people with similar views, values, beliefs, practices, attitudes, interests, and preferences, as well as cognitive distortions and unfair cultural biases, where diverse, outgroup voices tend to be excluded, discredited or suppressed. Such groups and contexts are thus characterized by reduced exposure to both heterogeneous information and relationships – a systemic reduction that is difficult to experience or implement at a large scale in environments that are not governed by algorithms. It follows that these AI systems undermine rather than expand the spreading of novel or diverse ideas and viewpoints from in-group perspectives, as well as of critical, unexpected, challenging relational encounters. In this sense, they can negatively affect both the socio-epistemic conditions (such as the ne-

cessity of exposure to diverse moral ideas and social practices and the development of a culture of justification) and the moral one *stricto sensu* (i.e., the widespread exercise of critical, open-ended moral reasoning) argued to be necessary for moral progress at the collective level, raising controversial implications for both individuals, as members of specific groups, and for groups themselves, as collectives.

Indeed, in groups of like-minded people, with similar ways of thinking and acting, decision-making biases included, people tend to radicalize, rather than mitigate by mutual compensation, their own cognitive errors and cultural biases, as well as their ideas, viewpoints, and attitudes, either by social pressure dynamics or because particular cognitive bias (such as confirmation bias). As a result, in-group individuals could end up developing increasingly self-enclosed identities (Parsell 2008) with increasingly rigid or impermeable physiognomy (Tiribelli 2023), making them less able to understand and respect those who have different ideas and especially outgroup perspectives, ultimately leading to increased social cleavage and division (Parsell 2008, p. 43).

Additionally, the ingroup social pressure and emotional tendency to uncritically adhere to dominant positions or to confirm what aligns with and reinforces our pre-existing beliefs or attitudes can lead to reduced reflective scrutiny, critical thinking, and moral reasoning. In other words, it can lead to a decreased demand for – and, over time, a gradual erosion of – our exercise of moral capacities, from the ability to develop genuine moral reasons to that of both demanding and providing genuine reasons (i.e., moral accountability and answerability).

In sum, these systems, by placing individuals in environments with reduced heterogeneous informational and relational exposure, preclude them by design from encounters with different social and moral ideas, beliefs, concepts, and practices, so disabling the exercise of moral reasoning. Over time, these phenomena can raise the possibility of generating a moral reasoning deskilling, that is, a weakening of our critical ability to assess critically whether what we are adopting and following from a moral standpoint (in terms of beliefs, concepts, rules, and practices) is optimal or requires revision or change – an ability that is largely understood as critical for moral progress in the related debate.

Similar implications can be extended to groups as collectives. Indeed, it is likely that groups with increasingly self-enclosed and radicalized identities, when called to engage in critical dialogues in the public spaces of our liberal-democratic societies, will be less open to mutual understanding and agreement, as well as less inclined to cooperation in inclusive terms. They may struggle to develop shared commitments and social projects (Giovanola & Sala 2021) and hence to act together towards common goals for a better society. Rather, hard group physiogno-

mies and the lack of ingroup widely deliberated and strongly and deeply felt genuine reasons to support group identities might lead groups to become intolerantly unreasonable with each other, creating *ecological conditions* for **moral regress**.

It follows that the possibility for the exercise of open-ended, critical moral reasoning is endangered not only at the individual level but also at the collective one, as is the possibility of moral progress.

4. Conclusive remarks: Moral progress by AI design?

In this article, we have shown how AI systems governing our pervasive digital ICTs can affect the necessary conditions for moral progress by reshaping our hyper-connected informational and socio-relational environments in ways that, instead of promoting, erode the exercise of moral reasoning of both groups, as collectives, and its members, as individuals, paving the way to the possibility of losing the moral gains we care most and of incurring to phenomena of moral regress. By doing so, we have disclosed not only the relevance of such an ethical inquiry in the light of the implications it outlines; we have also shown the need of bridging the scholarship on moral progress and that of AI ethics to act to ensure AI systems become a driver instead of a hindrance of moral progress.

This necessity is even more evident if we consider prominent practical solutions aimed at promoting moral progress and avoiding moral regress elaborated in the debate on moral progress within the field of moral philosophy.

Indeed, those who contend a narrow conception of moral progress tend to adopt proactive or intentional approaches to moral progress, grounded on the possibility of eliciting moral progress or guiding societies toward moral progress, through *design* actions. Many scholars of moral progress claim that the philosophical and empirical studies on moral progress can provide precious knowledge to shape our social environments in ways that promote instead of hampering moral progress (Buchanan 2021, p. 236). However, the proposals made thus far in this debate focus only on institutional design and are mainly oriented by criteria of economic or political efficiency. In this regard, for example, Buchanan (2021) has criticized the lack of practical solutions that consider the moral effects of diverse typologies of institutional design.

Based on the ethical inquiry we sketched out in this article, we claim here that there is another gap in this regard that we might fill in to create favorable ecological conditions for moral progress.

If champions of moral progress invite to reshape or design social environments by considering knowledge on what promotes or hinders cer-

tain preconditions required by moral progress, we claim here we cannot prescind to also consider that (i) AI systems reshape such preconditions and therefore (ii) can translate into forces *for* or *against* moral progress. Thus, any design actions for promoting moral progress *should* seriously consider how intentionally re-designing interconnected AI systems considering literature on moral progress so as to make them conducive instead of adverse to it.

To sum up, in this article we hope to have shed light on the need to ethically design such systems, which means, in this case, developing design criteria that consider, in a novel way, studies on moral progress. Indeed, if such technologies show, theoretically, the potential to adversely affect moral progress, the same if redesigned considering both the *sine qua non* conditions of moral progress and the moral risks and effects that the action of AI systems on such conditions can generate hold the potential to become not only instruments of protection from moral regress, but also real forces of concrete promotion of moral progress.

Bibliography

Benjamin, R.,
2019 *Race after technology: abolitionist tools for the new Jim code*. Polity, Cambridge.
Buchanan, A.,
2020 *Our Moral Fate: Evolution and Escape from Tribalism*, The MIT Press, Cambridge.
Buchanan, A., Powell, R.
2018 *The Evolution of Moral Progress. A Biocultural Theory*, Oxford University Press, New York.
Fricker, M.,
2007 *Epistemic injustice: Power and the ethics of knowing*, Oxford University Press, Oxford.
Giovanola, B., Sala, R.,
2021 *The reasons of the unreasonable: Is political liberalism still an option?*, "Philosophy & Social Criticism",1226-1246.
Giovanola, B., Tiribelli, S.,
2022 *Weapons of moral construction? On the value of fairness in algorithmic decision-making*, in "Ethics and Information Technology", 24, 1-13.
Giubilini, A., Savulescu, J.,
2018 *The Artificial Moral Advisor. The "Ideal Observer" Meets Artificial Intelligence*, in "Philosophy & Technology", 31(2), pp. 169-188.
Horkheimer, M.,
1937 *Traditional and Critical Theory*, The Continuum Publishing Company, New York.
Pariser, E.,
2011 *The Filter Bubble*, Penguin, London.

Parsell, M.,
2008 *Pernicious virtual communities: Identity, polarization and the web 2.0*, in «Ethics and Information Technology».
Jaeggi, R.,
2018 *Critique of Forms of Life,* Harvard University Press, Cambridge.
O'Neil, C.,
2016 *Weapons of math destruction: how big data increases inequality and threatens democracy*, Crown, New York.
Nguyen, C.T.,
2020 *Echo Chambers and Epistemic Bubbles*, in "Episteme", 17(2), 141-161.
Sauer, H., Blunden, C., Eriksen, C., Rehren, P.,
2021 *Moral progress: Recent developments*, in "Philosophy Compass", pp. 1-10.
Savulescu, J., Maslen, H.,
2015 *Moral Enhancement and Artificial Intelligence: Moral AI?*. In: Romportl, J., Zackova, E., Kelemen, J. (eds.) *Beyond Artificial Intelligence. Topics in Intelligent Engineering and Informatics*, vol 9. Springer, Cham.
Singer, P.,
(1981) 2011 *The expanding circle-Ethics, evolution, and moral progress*, Princeton University Press, Princeton.
Sunstein, C.,
2008 *Democracy and the internet*. In: van den Hoven, J., Weckert, J. (eds.) *Information technology and moral philosophy*. Cambridge University Press, New York, pp 93-110.
Tiribelli, S.,
2022 *Moral freedom in the age of artificial intelligence*. Mimesis International, Milan-London.
Tiribelli, S.,
2023 *Inequalities and Artificial Intelligence*, in "Filosofia Morale/Moral Philosophy", (3).
Tiribelli, S.,
2023 *Identità personale e algoritmi. Una questione di filosofia morale*, Carocci editore, Rome.
Yeung, K.,
2018 *Algorithmic regulation. A critical interrogation*, in "Regulation & Governance", pp. 505-523.
Zuboff, S.,
2019, *The age of surveillance capitalism: the fight for a human future at the new frontier of power*, Public Affairs, New York.