



Critical Review of International Social and Political Philosophy

ISSN: 1369-8230 (Print) 1743-8772 (Online) Journal homepage: www.tandfonline.com/journals/fcri20

The future of work: freedom, justice and capital in the age of artificial intelligence

Filippo Santoni de Sio, Txai Almeida & Jeroen van den Hoven

To cite this article: Filippo Santoni de Sio, Txai Almeida & Jeroen van den Hoven (2024) The future of work: freedom, justice and capital in the age of artificial intelligence, *Critical Review of International Social and Political Philosophy*, 27:5, 659-683, DOI: [10.1080/13698230.2021.2008204](https://doi.org/10.1080/13698230.2021.2008204)

To link to this article: <https://doi.org/10.1080/13698230.2021.2008204>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 13 Dec 2021.



[Submit your article to this journal](#)



Article views: 14125



[View related articles](#)



[View Crossmark data](#)



[Citing articles: 35 View citing articles](#)



OPEN ACCESS



The future of work: freedom, justice and capital in the age of artificial intelligence

Filippo Santoni de Sio, Txai Almeida and Jeroen van den Hoven 

ABSTRACT

Artificial Intelligence (AI) is predicted to have a deep impact on the future of work and employment. The paper outlines a normative framework to understand and protect human freedom and justice in this transition. The proposed framework is based on four main ideas: going beyond the idea of a Basic Income to compensate the losers in the transition towards AI-driven work, towards a *Responsible Innovation* approach, in which the development of AI technologies is governed by an inclusive and deliberate societal judgment; going beyond a philosophical conceptualisation of social justice only focused on the distribution of ‘primary goods’, towards one focused on the different goals, values, and virtues of various social practices (Walzer’s ‘spheres of justice’) and the different individual capabilities of persons (Sen’s ‘capabilities’); going beyond a classical understanding of capital, towards one explicitly including mental capacities as a source of value for AI-driven activities. In an effort to promote an interdisciplinary approach, the paper combines political and economic theories of freedom, justice and capital with recent approaches in applied ethics of technology, and starts applying its normative framework to some concrete example of AI-based systems: healthcare robotics, ‘citizen science’, social media and platform economy.

KEYWORDS Ethics of artificial intelligence (AI); future of work; Sen’s capability approach and technology; Walzer’s spheres of justice and technology; technological unemployment and universal basic income; Responsible Innovation

Introduction

Artificial Intelligence (AI) is a broad term covering any technique designed to solve problems traditionally assigned to human intelligence (Amoroso & Tamburrini, 2019). Recent developments in AI have raised the question of its possible impact on the level and quality of employment (Brynjolfsson & MacAfee, 2014; Ford, 2015; Frey & Osborne, 2017; Willcocks, 2020). This paper addresses the question, which theories of freedom, justice and capital are needed in order to ensure that AI will have a beneficial impact on human work. Whereas freedom, justice, and capital are concepts grounded in political philosophy and economy, the broader issue of the impact of technology

CONTACT Filippo Santoni de Sio  f.santonidesio@tudelft.nl

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

on human values has been more systematically addressed in philosophy of technology (Scharff & Dusek, 2014) (Van Den Hoven et al., 2015). Therefore, in our analysis we discuss and combine political theories of freedom and justice (Pettit, 2008; Rawls, 1971; Sen, 2009; Van Parijs, 1995; Walzer, 1983a) and theories in applied ethics of technology (Calo & Rosenblat, 2017; Santoni de Sio & van Wynsberghe, 2016; Stilgoe et al., 2013; Van Den Hoven, 2013). One methodological goal of this paper is indeed to promote a closer collaboration between these disciplines on this topic. On a more substantive level, this paper aims at providing a *general theoretical normative framework* to analyse and govern the potential impact of AI on freedom and justice in relation to work. We identify three theoretical obstacles to the creation of such a framework, and we correspondingly start identifying three ways to overcome them.

The first obstacle is the virtual *absence of conscious choice* in decisions that affect the direction of technology and innovation, and the lack of a theoretical framework that could guide us in deciding how technological developments and innovations should be judged and guided. 'By default', technological development is determined by the conventional theoretical and practical factors that drive societal developments today: (cost-)efficiency and productivity as pushed by big economic players. That needs reconsideration. Rather than merely proposing to compensate the losers in the transition to AI-based work via a Universal Basic Income (UBI) (Van Parijs & Vanderborght, 2019) (Walker, 2014), we advocate to follow a *Responsible Innovation* approach, in which decisions regarding the design, development, introduction, and governance of AI technologies are reached through a deliberate societal judgment explicitly governed by a broader set of individual and societal values (Stilgoe et al., 2013; Van Den Hoven, 2013). As we argue, this more pro-active approach can better support the promotion of the political freedom cherished by supporters of UBI such as 'real libertarians' (Van Parijs, 1995) and neo-republicans (Casassas, 2008; Lovett, 2018; Pettit, 2008).

The second, related, obstacle is a philosophical conceptualisation of social justice too focused on the distribution of (material) resources. We invite to go beyond a broadly Rawlsian approach to social justice based on the reduction of (unjustified) inequities via a fairer distribution of wealth or other 'primary goods' (Rawls, 1971). We propose to look at two alternative approaches, both critical of Rawls approach, which, in different but complementary ways, point to the context-specific circumstances under which technology may promote people's freedom and wellbeing. The first is Michael Walzer's idea of 'spheres of justice', according to which social justice can be achieved only by preserving the specific goals, values and virtues of different social practices (Walzer, 1983a). The second is the idea of the development of individual human capacities as the ultimate end of political action, which can be found in Amartya Sen's 'capability approach' (Sen, 2009).

The third and last obstacle is the lack of (theoretical) economic underpinnings for a society that is able to govern the development and use of technology in a just way. In currently dominant economic theory, the role of mental capacities and activities is not sufficiently recognised as a crucial part of capital, and mental capacities are therefore not sufficiently recognised, rewarded, and supported (Naastepad & Mulder, 2018). This paper criticises the ‘myth of autonomy’ of digital technology and propose to recognise human mental capacities as a capital to be rewarded.

The next sections present these three proposals and their general implication for the design, use or regulation of AI-based systems. In line with our methodological goal of combining political philosophy and applied ethics of technology, each theoretical section is accompanied by a concrete example of a design of a technological practice which may realise the theoretical approach. Applications discussed in the paper are healthcare robotics, ‘citizen science’, social media and platform economy.

Human freedom from basic income to responsible innovation

Universal Basic income (UBI)

The AI revolution may have a deeper and longer-term impact on the future of work than previous industrial revolutions (Brynjolfsson & MacAfee, 2014; Ford, 2015). AI allows to automate not only repetitive physical tasks, but potentially any task that can be decomposed, including more complex tasks like driving as well as many mental activities that had been for long thought to be not open to automation, like accounting, planning, doing legal research or medical diagnosing (Ford, 2015). Moreover, the digital market has shown a consistent tendency to take a ‘winner take all’ structure (Frank, 1996). We might be eventually facing what John Maynard Keynes (Keynes, 1930) called ‘technological unemployment’ (Naastepad & Mulder, 2018). The extent to which AI will lead to a (big) net job loss or rather to a (big) job restructuring is very debated. In particular, (Willcocks, 2020)’s analysis of the more recent studies suggests that a big net job loss is unlikely, and that ‘the much bigger storyline is of skill disruption and change’ (Willcocks, 2020, p. 297). He also stresses that future research needs to address the ‘important policy, work design and social responsibility implications’ of these changes (p. 297).

In fact, various proposal have already been advocated for, to prevent or mitigate the impact of AI on the life of people, especially those in precarious positions at the receiving end of the digital economy: a stronger emphasis on social and worker rights (Stewart & Stanford, 2017), the need for a reinvigoration of unions (Hacker & Pierson, 2010), property owning democracy in relation to data (Loi et al., 2020), stronger policies for the redistribution of income, and various versions of robot taxes (Abbott & Bogenschneider, 2018).

One particularly radical proposal is the introduction of Universal Basic Income (UBI) (Ford, 2015; Walker, 2014), a policy that would periodically provide each citizen a fixed amount of money, irrespective of their income from other sources or willingness to work.

UBI has a long history and has been argued, in different forms, by thinkers with very different ideological profiles such as Thomas Paine (1797), Martin Luther King (1967), and Milton Friedman (1962). The concept entered the contemporary political debate with the work of Philippe Van Parijs, synthesized in his book *Real freedom for all* (Van Parijs, 1995). Van Parijs describes himself as a 'real libertarian'. He believes that self-ownership (the paramount value of libertarianism) can be achieved only if people have an unconditional access to the resources needed to pursue their own conception of the good life. Thus, he proposes the provision of UBI 'at the highest sustainable level' (Van Parijs, 1995, p. 5).

After Van Parijs, the idea of UBI continues to be taken up by many different, sometimes conflicting, ideological perspectives: a feminist (Alstott, 2001; Elgarte, 2008), a socialist (Calnitsky, 2017; Wright, 2006), a libertarian perspective (Zwolinski, 2019). One interesting alternative to Van Parijs' idea of UBI as a means to promote 'real freedom' is offered by so-called neo-republican thinkers (Casassas, 2008; Lovett, 2018; Pettit, 2008). Neo-republicans reject the classical liberal notion of freedom, which holds that people are free so long as their choices are not interfered with by other agents. According to the neo-republican notion of freedom, also called non-domination, people are free so long as no agent is in a position of power to interfere with their choices. Thus, whereas liberals identify unfreedom as actual choice-interference, neo-republicans identify unfreedom (or domination) as potential choice-interference. As for domination at the workplace, some neo-republicans believe that UBI may mitigate the power imbalance between employers and workers, by making the latter less dependent on the former. The idea is that the financial resources provided by UBI give workers a real ability, as opposed to a merely a formal right, to exit or avoid entering jobs that are exploitative or otherwise undesirable (Casassas, 2008; Lovett, 2018; Pettit, 2008).

So, according to two popular arguments, UBI enhances people's freedom. It allows them to pursue their preferred conception of the good life (Van Parijs' point) and ensures that workers are not in a position of being dominated by their employers (the neo-republican point).

Responsible Innovation (RI)

The above-mentioned arguments have received different criticisms (Bidadanure, 2019) typically UBI being unfair (Van der Veen, 1998; Van Donselaar, 2009) and unfeasible (Gourevitch & Stanczyk, 2018). Be that as it

may, in this paper we raise a different criticism: UBI seems to accept technological unemployment as a destiny, and this arguably undermines the aims of both Van Parijs' real libertarianism and neo-republican non-domination.

First of all, at least in the eyes of those who propose it as an answer to the risks of technological unemployment caused by Artificial Intelligence (e.g. Walker, 2014) – UBI is a *post-hoc* intervention needed to compensate those whom the digital revolution will (permanently) leave without a paid work. So, these basic income strategies seem to accept technological unemployment as a destiny, the inevitable consequence of technological development; or at least they are silent about possible strategies to *prevent* this structural mass unemployment to occur in the first place. This is a problem for 'real freedom'. Real freedom requires not only that people are not obliged to do paid work they do not consider as fair or valuable but also that they are able to do work that they consider as valuable. Whereas 'the end of work' maybe good news for those who see it mainly as a source of suffering and exploitation (Danaher, 2017), (paid) work is for some a source of many 'goods other than money' (Gheaus & Herzog, 2016): acquisition of skills and self-identity, engagement in social life, and others. Moreover, one may want to have a paid job to avoid the stigma and discrimination associated with unemployment in industrialised society (Sage & Diamond, 2017). So, UBI risks to become in this perspective more of a monetary compensation for the loss of freedom of choice than a means to protect or increase it.

Also, some UBI supporters believe that UBI will make the persons left unemployed by technological change willing and able to engage in other (more) valuable forms of work activities outside the boundaries of traditional paid contracts. However, it is unclear where they might take the resources and develop the skills and motivation to do so. UBI and other forms of financial support would in the best-case scenario be sufficient to provide the means to make one's living, not the resources to develop the capacities and opportunities to engage in new social, intellectual, professional, leisure or artistic activities. In fact, many doubt that UBI might be ever be high enough to grant workers real independence from the labour market (Birnbbaum & De Wispelaere, 2020) so that it sounds implausible to assume an UBI sufficiently high to cover also for investments in self-development, education, meetings, travels etc. Moreover, personal self-development requires not only material resources but also a social environment that promotes and supports it (Sen, 2009).

Accepting technological unemployment is also a problem from the neo-republican point of view. First of all, neo-republican theories of UBI rely on the idea that financial independence granted by UBI will give employees more power to walk away from dominating work relationships. One objection to this idea is that workers may have internalised the intimidation of their bosses and/or not being confident enough to take the decision to leave. Gourevitch

(2013, p. 608). They may also reasonably believe that they are unlikely to find a new job which is not similarly dominating, insofar as it is a ubiquitous practice among capitalist enterprises to give managers sufficient power to make their workers' life miserable or even humiliating (Anderson, 2017). This points to a general limit of the liberal and neo-republican approaches to freedom and domination, namely that they see domination as something personal, happening between individual persons. Unlike structural accounts of domination (Foucault, 1955; Young, 1990) neo-republican accounts fail to recognize social, economic or technological structures as potential sources of domination (Gourevitch, 2013; Thompson, 2013).

To be sure, individual employers do play an important role here. As put by Martin Ford, the rapid development of digital and autonomous technologies carries the risk of creating mass unemployment or worsening of the quality of working conditions, also insofar as this is very much welcomed by 'rational business owners' who 'despite all the rhetoric about "job creators," [...] do not want to hire more workers' (Ford, 2015, pp. 255–6; Noble, 2011) or to invest in the improvement of their employees' quality of life. However, as we argue, for this risk to materialise, an equally important condition is that the research, design, development and governance of such technologies is driven and accompanied by the same general political and economic ideology dominated by the values of cost-saving and increase in productivity. Therefore, one crucial component is acting on this ideology, rather than only trying to compensate for the damages it may create.¹

In contrast with the basic income strategies – both the real libertarian and the neo-republican one – our claim is that freedom can be better protected by the systematic application of the ideal known as *Responsible Innovation* to the design, development, governance and use of Artificial Intelligence. According to Responsible Innovation, no matter how formidable technology and the economic forces propelling it is, we should not succumb to an (implicit) technological determinism that obscures the institutional and socio-technical design space and discourages to propose new designs. In fact, technological development can be beneficial to society only if designers and policy makers are willing and able to pro-actively embed more relevant human and social values into the process of design and development of the new technological systems, as opposed to just regulate the use of technology or devise post-hoc policies to govern the ethical and societal impact of technology (Van Den Hoven, 2013).

The basic ideas behind the Responsible Innovation can be summarised as follows: a) technological development is always informed by specific values of specific stakeholders – technology do 'have a politics' (Winner, 1980) – and therefore b) the time in which value choices and policy interventions can still make a difference are the early stages of design and development of technology. Once this process of design and development has been deeply

shaped by certain values and certain stakeholders' interests, it is far more difficult if not impossible to see to it that technology will serve other values and interests (Van Den Hoven, 2013). More generally, RI requires that the innovation process is made more transparent, inclusive, and responsive to the interest and values of stakeholders and the public (Stilgoe et al., 2013).²

Therefore, adopting the Responsible Innovation approach requires to make an important change to the basic moral question as framed in the debate on UBI. This debate revolves around the question: To what extent, in a future in which technology has worsened the conditions of workers, we as a society can legitimately incentivise or even force those (few) who have dramatically benefitted from a cost-efficiency-driven technological development to give some financial compensation to those many whose life conditions have been dramatically damaged or impoverished by it. The question should rather be: to what extent we as a society should allow and even embrace such a cost-efficiency-driven technological development with its related existential and social effects in the first place. Therefore, Responsible Innovation requires, first of all, that we embrace a deep change in our theoretical understanding of the current technological and economic trend, and that we strive for a more inclusive technological and innovation process.

How to realise the ideal of Responsible Innovation in the development of digital technologies which may improve rather than worsen the conditions of workers is a broad question to be addressed from many different angles, including psychology, social science, and engineering. In this paper we focus on one specific element: a better philosophical understanding of the concepts of freedom, justice and capital to be embedded in the design, development and governance of Artificial Intelligence. We propose to introduce in the philosophical debate on the future of work views of justice that go beyond the dominant approach based on wealth and goods distribution; and a theory of capital which goes beyond its, still dominant, classical understanding as the sum of material and financial resources. We will present these theories of justice and capital in the next sections. Each theoretical section is complemented by a practical section, proposing examples of possible application of the theory to the development of AI-technology.

Justice from theory to practice

The philosophical debate on social justice of the last decades of the XX century has been largely dominated by John Rawls' *Theory of Justice* (Rawls, 1971). Broadly speaking, this framework sees equality mainly as an ideal concerning the distribution of resources. – i.e. people are seen as entitled to an equal (or leximin) share of society's distribution of resources.³ Relevant resources whose distribution should be regulated by a theory of justice are called by Rawls 'primary goods' – that is 'all-purpose means' such as income,

wealth, 'the social basis of self-respect'. Famously, according to Rawls' 'difference principle', differences in the distribution of primary goods are ethically acceptable only if they raise the level of the least advantaged in society (Rawls, 1971). The founding father of the movement for Universal Basic Income, Philippe van Parijs, claimed that in order to expand as much as possible the set of 'all-purpose means' available to the least advantaged, we should establish an unconditional basic income at the highest sustainable level (Van Parijs, 1992, p. 21).

However, successful as it has been in the renaissance of normative political philosophy, Rawls' approach has also been criticised for its being 'too detached from reality to guide political action' (Valentini, 2012, p. 654), in particular for its being designed to work only within national States (Pogge, 2008), for its presupposing agents' compliance with its demands while obfuscating the importance of avoiding power relations which make compliance with the demands of justice difficult to realise (Pettit, 1996, 2012), for its focus on fair procedures rather than on their material effects on real persons (Nussbaum, 2011; Sen, 2009).

To be fair, Rawls is also interested in the stability of real political systems, and he emphasises that, as a matter of facts, his principles of justice are likely to support more stable social arrangements than utilitarian ones by being grounded in a moral psychology of reciprocity rather than in one of sympathy (Rawls, 1971, Chapter 7). However, as we take the above criticisms to remain substantially fitting in the case of the just design of AI, we propose to go beyond Rawlsian distributive justice and to ground future visions on the design and implementation of these technologies in two alternative approaches to social justice: Michael Walzer's theory of justice as the 'art of separation' between 'spheres of justice' and Amartya Sen's 'capability approach'.

Spheres of justice

In *Liberalism and the art of separation* Michael Walzer complained that classical liberal theorists failed to see the political weight of the social forces of individual wealth and corporate power, and that radical inequalities of wealth often generate power relationships which make some exchanges only formally free (Walzer, 1984, pp. 321–322). Therefore, in order to realise a society that is really free, that is one in which such kinds of coercion and domination are ineffective, a 'positive structure' is needed – a complex set of 'institutions, rules, mores, and customary practices' (322).⁴ In other words, the liberal art of separation can work only if it is accompanied by 'disestablishment and divestment' of economic powers, and if 'appropriate cultural forms develop within the economic sphere' (322). In this perspective, distributive justice is not just a matter of transferring wealth or other general resources from the

wealthy to the poor; it is something different and more complex – in Walzer words, is a ‘matter of getting the lines right’ (323) and building political and institutional – we will add: socio-technological – structures to prevent unwanted influences between the political, the civil and the economic domains.

In *Spheres of Justice* (Walzer, 1983b), Walzer further explains how these lines should be drawn. Whereas traditional liberals have mostly drawn the lines around individuals’ activities and rights, Walzer proposes to look at the principles defining and governing different social domains and activities, which he calls ‘spheres’. According to Walzer, different societal spheres revolve around different ‘goods’ – in a moral/social not economic sense – and are defined by different constitutive aims and purposes. Health care is about medical care, politics is about public office, the market about exchanging material goods, education about learning etc. These different spheres also have their own criteria, norms and moral rules that govern the regulation, governance and distribution of things that matter. Heart surgery is, for instance, usually given on the basis of need to those who need it most, public office to those who were democratically elected, money by means of some market mechanism for value added and work done, etc. According to Walzer, *pace* Rawls and even more Van Parijs, there are no universal principles and ‘all-purpose’ primary goods that can validly be applied across all spheres of life, and it is widely considered inappropriate to use criteria and normative rules from one societal sphere to guide behaviour and decision making and distribution in another sphere. Were that to happen, things would get mixed up that ought to not be mixed up. In the market it may be alright for example, to compromise, bargain, strike deals, cut corners or go for the quick win. In science, politics, education, and family life our choices and decisions should arguably not be affected by this market logic rife with compromise and bargains. By the same reasoning profit motives are considered inappropriate in health care. This is not the same as saying that cost and efficiency considerations have no role to play in health care or other sectors. A just society according to Walzer is one that exhibits ‘complex equality’, i.e. a just society respect the different moral logics of different social spheres, roles, tasks, aims.

The art of separation consists therefore in avoiding normative cross contamination, or moral short circuits. For similar reasons we reject the idea that people can convert advantages in one sphere to advantages in another sphere, and the idea that possession of one type of good (e.g. health, money, education) will also get you everything else: public office, university degrees, places on a waiting list, and love. This idea of social differentiation in systems, subsystems, system boundaries and the relative autonomy of these subsystems can also be found in work of sociologists such as Luhmann and Bourdieu. Moreover, Walzer’s proposal to protect different spheres of justice resonate well with Jürgen Habermas’ criticism of modern capitalism as

leading to the 'colonisation of the lifeworld' (Habermas, 1987). In Habermas' terms, money and power have colonised areas of social life that are dependent on communicative actions, that is on social relations driven by different normative expectations (Jütten, 2011). Many have formulated generic ethical principles for the use of AI and robots in society (Floridi et al., 2018), construed as homogenous devoid of relevant local meanings that pertain to different social spheres, and that need to be accommodated by the ethical principles. As we will see in the next section, Walzer's perspective provides us with justified constraints and shared moral requirements for design of future technological systems. The next section also highlights a potential connection between the idea of design AI for 'spheres of justice' and another important philosophical tradition: virtue ethics.

Application #1: healthcare robotics

The Responsible Innovation approach, coupled with Walzer's idea of spheres of justice, can help framing the issue of automation and the future of work in an original normative perspective. Two new questions become relevant here: first, which aspects of which spheres of life and work should be delegated to technology; which in turns depends, as we argue, on how automation will affect the achievement of relevant and shared moral aims, the quality of the process of achieving those aims and the moral opportunity cost on where automation contributes to furthering and achieving the constitutive aims of those societal spheres; second, how to *design* technologies and institutions which promote the achievements of the values and goals of specific practices. Take, as one among many possible examples, the use of robotic systems in healthcare. Some studies suggest that progress in robotics can also be one possible answer to the issue of aging in Europe: since according to demographic trends humans might not be enough to cover for healthcare tasks, human care workers can be assisted and even replaced by care robots in their job; at least in those tasks which can be performed also by intelligent machines (Sharkey & Sharkey, 2012). However, in this perspective the design, regulation and use of healthcare robotics should keep into account the different dimensions of healthcare practices and their different constitutive aims. Healthcare is a broad concept, which includes both very technical tasks like, for instance, those included in precision surgery, whose main goal is the effective realisation of a certain material outcome (e.g. the successful replacement of a cardiac valve), as well as activities with a broader and less material goal like improving the persons' well-being.

As an example of how the principles of Responsible Innovation and the logic of the separation of spheres of justice may shape the design and introduction of robots in the healthcare domain, Aimee van Wynsberghe

and Filippo Santoni de Sio (Santoni de Sio & van Wynsberghe, 2016; Van Wynsberghe, 2015), have analysed an apparently simple and straightforward task: the collection of urine sample for testing the presence (or lack thereof) of chemotherapy toxins for children undergoing chemotherapy. From the point of view of effectiveness and (cost)efficiency, the collection of the sample may arguably be better performed by a robot; moreover, the sample collection may be embarrassing for the patient and dangerous for nurses' health, and removing the human from the task would prevent these further unwanted risks. However, this task is also part of a broader care activity, whose goal is the maintenance or restoration of the well-being of the patient; this seems to require, among other things, that the patient enjoys some interaction with a human care-taker, that the care-taker takes responsibility for the therapy, and that the patient maintains a relationship of trust with the human care-takers. Therefore, even a relatively simple process like the urine sample collection may not be fully automated and a human presence is required along the process, even though possibly not at the very moment of the collection itself. One may even go further and, following the care ethics tradition, claiming that one constitutive aim of a care practice is the realisation of the care skills of a (human) nurse, and this would be one additional reason to maintain human nurses involved in all sensitive care activities.⁵

Similar analyses can be made in relation to the educational, the judicial, the political, the private life and other spheres, in which productivity and efficiency are not the only and ultimate value at stake. One general normative consideration supporting the idea of design for 'spheres' is one deriving from the tradition of virtue ethics, according to which certain activities are worthwhile insofar as they realise and support certain attitudes and habits that are considered to be central to a flourishing human life (Vallor, 2016).

The general point is that by having the design, development and use of robots driven by the logic of sphere-appropriateness, it may happen that, as a matter of fact, more opportunities of work will remain available for humans. This is a very different perspective than the one prescribing to 'protect human work' as a goal in itself (Marchant et al., 2014) or inviting to find a way for persons to remain relevant in the new machine-driven world, whatever this will turn out to be (Brynjolfsson & MacAfee, 2014). Here the suggestion is that the prevention of technological unemployment might partly happen, as it were, as a side-effect of a politics and an economics of technology informed by the recognition of different values as constitutive of different activities; and which promote the design of technologies and institutions which reflect the appropriate values of different spheres of life.

The capability approach

A different but related criticism to Rawls' approach to justice came from Amartya Sen. According to Sen, distributive justice should focus not only on the amount of primary goods or resources allocated to different persons, but rather on what each person is actually able to do with the available goods and resources, given her personal skills and talents, in the natural and social context in which she finds herself to live. In Sen's words: 'justice cannot be indifferent to the lives that people can actually live' (Sen, 2009, p. 18). Unlike Walzer, Sen's focus is not much on the diversity in moral reasoning underlying values, goals (and virtues) of different domains. Rather, Sen argues that different circumstances also provide individual people with different ability to convert resources into levels of well-being.

Sen's approach can be another powerful conceptual tool to address the question of the fair design, regulation and use of AI, insofar as it helps give a new and concrete content to two traditional normative ideals: that technology (should) make people's life better, that is increasing people's well-being; and that technology (should) expand people's opportunities to realise their goals, that is people's freedom. As for well-being, Sen strongly criticizes the traditional utilitarian idea, which has also permeated important strands of economic theory, which equates well-being with the realisation of subjective 'utilities' and/or the satisfaction of unanalysed sets of preferences (Little, 2011; Sen, 2018). In fact, firstly, as a matter of fact, people's preferences and desires tend adapt to certain unfavorable circumstances and therefore the self-evaluation in terms of satisfaction or happiness may not be a good indicator of the persons' well-being.⁶ For instance, someone who has lived most of her life under an oppressive regime may simply be unable to appreciate the value of freedom and may even stop desiring it. But this does not necessarily mean that – other things being equal – she may not be better off in a more liberal one (Sen & Williams, 1982). To evaluate people's well-being in terms of capabilities guarantees a more objective picture of people's life. Moreover, reasoning in terms of capabilities also helps including the importance of human agency and choice in the definition of well-being. Sen believes that achieving some goods via the pursuing of free agency, that is while having the capacity of doing so is – other things being equal – morally better than just receiving the same good as the result of a concession by others, or possibly as a result of an unfree choice. In this respect the capability approach gets close to the neo-republican tradition, which insist on people's 'effective control' on their life as an independent dimension of evaluation of well-being (Qizilbash, 2016).

As for freedom, Sen has proposed a radical revision of its classical libertarian definition, equating it to the absence of external constraints. In this perspective are constraints on freedom only those obstacles that are brought

about by other agents, typically varieties of human coercion. Physical disabilities or natural circumstances like scarcity of food due to drought are not considered as a cause of unfreedom (Carter, 2008). According to this narrow definition of freedom, not only do natural constraints like physical disabilities or natural circumstances ended up not being considered as a cause of unfreedom, but even obstacles created by impersonal economic and social forces like a dramatic unequal distribution of resources, lack of education, or pervasive social discrimination couldn't be considered as factors limiting people's (political) freedom; therefore, in this view these are not circumstances about which policy-makers can and should do actively anything about either.

We have argued above that both of these assumptions seem to be unwarranted. On the one hand, the fact that disabilities rooted in physical and natural conditions can be best dealt with by engineering or medical interventions does not entail that these circumstances do not *also* call for ethical consideration and political intervention. In fact, one crucial presupposition of the Responsible Innovation approach that we endorse is that science and engineering are also political enterprises crucially shaped by value choices, and that should therefore be subject to appropriate political and economic governance. From this same perspective, the fact that no individual may easily be blamed or considered as the direct cause of a complex social, economic – and we will add: technological – setup which negatively affect the life of many people, doesn't seem in any way to lift policy-makers from the responsibility to intervene to modify this state of affairs. Sen's focus of human capabilities – broadly conceived – as the central element of well-being and freedom nicely captures and summarizes all these elements. We therefore propose that it should also be a pillar of any policy for the development of Artificial Intelligence which aims to respect social justice and promote freedom and human well-being.

Application #2: citizen science

The Capability Approach has already been recently proposed as a framework to ground and systematize many ideas already present in design theory, such as 'participatory design' and 'inclusive/universal design' (Oosterlaken, 2013). It has also been much utilized in the field of so-called ICT for Development (ICT4D). In this domain, (Frediani & Boano, 2012), for example, have proposed that we apply the capability approach both to the process of designing technologies and to its outcome, i.e. the resulting technical systems; that is that we promote a broader participation in technological process, for instance, by adopting some form of 'participatory design', while at the same time making sure that the final outcomes are also supportive of the

relevant human capacities. This same idea can be applied to the development of non-professional participation in the design and realization of scientific research, also known as citizen science.

In the digital or 'p2p economy' it is possible to recognise, encourage and reward not only the individual people capabilities to produce economic values via digital platforms (Bauwens, 2012), but also the individual capabilities to contribute to the production of scientific knowledge. Diana Wildschut reports how, thanks to the recent advancements in computer and sensor technology and the creation of cheap open source micro-controllers boards, individual lay persons are now able to measure 'whatever they can think of in their environment' (Wildschut, 2017). They can log all their data and compare them with those of other people's. Online communities have been created around the hardware, but also around research topics, like spectralworkbench.org, where the spectra taken of different light sources are shared.

In this respect, suggests Wildschut, we seem to be back to the early days of the scientific revolution, where 'any interested person with time to spare could look at the stars, gather data and make predictions', and 'scientists were often also farmers, artists or statesmen.' To be sure, we do not want to go back to the pre-scientific revolution age – but it may be worth to exploit technological innovation to involve also more non-professional scientists in the production of scientific knowledge. However, in many current citizen science projects run by academic scientists, participants typically provide their time or computer power: the participant role is limited to data collection or classification. The participant's intelligence or creativity is not sufficiently utilised, the project only offers very limited opportunities for personal growth (Rotman et al., 2012), especially insofar as participants are rarely involved in the design of the research. This is arguably a missed opportunity in many respects. From a science perspective, the design of research projects may be improved by including citizens' local knowledge and local networks (Guimarães Pereira & Saltelli, 2017; Pellow & Fischer, 2002; Warren, 1991). From a policy and engineering perspective, local governments may fruitfully use the input of their citizens to govern cities to their citizens, and citizens communities may come up with small scale solutions for energy or local food production, or for the recycling of waste. From a political perspective citizen science may sometimes be an antidote to political and institutional corruption as nicely shown, for example, by the Flint water case in Michigan (Barrie-Jester, 2016).

Citizen science is one example of the way in which technological innovation coupled with new appropriate institutional design and governance may contribute to the development of people capabilities, rather than to their replacement. To be sure, the extent to which citizen science and other platform-based social activities may really promote freedom, agency and human capabilities, crucially depends on the way these platforms and activities will

be organised. (Mirowski, 2017) rightfully insists that to produce real knowledge citizen scientists have to undergo a real training, 'not a drive-by blast of methodology lite'. In the terms of this article, meaningful agency requires capabilities. On a different and broader note, Mirowski notices that the majority of existing citizen science consists of the public donating its unpaid work and data to privately owned, online entities that subsequently digest it as 'big data'. He also warns that citizen science may turn to be yet another tool for the public sector to reduce their investments in science. The broad concern that digital platforms may be a tool of exploitation of human work rather than emancipation is addressed in the next section.

Human intelligence as capital

Ro Naastepad and Jesse Mulder (Naastepad & Mulder, 2018) have suggested that among the factors which stand in the way of a more responsible innovation and entrepreneurship in the digital economy is also the philosophical foundation of conventional economic itself. Not only because of its exclusive emphasis on labour-saving technological progress, but also because of 'its curious disregard of the further development of human capacities, in particular creative mental capacities' (Naastepad & Mulder, 2018, p. 4). Current economics is still 'informed by a narrow [...] conception of progress that focuses on material welfare and does not see the value of a free mental life' (Naastepad & Mulder, 2018, p. 13). Therefore, as it happens, the money formed as a result of productivity growth usually ends up on the balance sheet of firms as capital and, in today's economy, most of this capital flows to financial markets. In a nutshell: Human mental capacities are a crucial factor in the creation of a productivity growth, but the surplus deriving from this growth is only in minimal part invested in them.

Naastepad's and Mulder's reasoning may help to capture, among other things, the neglected role of human intelligent activities in the digital economy. Companies running search engines like Google, media platforms like YouTube and social networks like Facebook certainly derive their wealth from the intelligence and creativity of their managers, designers and developers, as well from the quality of their technologies. But these rely in turn on the scientific research and work of countless others, as well as from the constant activities of individual users who provide them with new data and information, which are in turn the result of their individual intelligent and creative activities: the research they perform, the video they produce, the picture they shot, the texts they create, the personal experiences they share. Giant digital firms can have a relatively low number of employees compared to traditional firms, not because the employees' work have been taken over by artificial intelligence, but rather because this has been taken over, and performed, usually for free, by its billions of users. In this respect it has recently talked of

digital labour (Fuchs & Sevignani, 2013). In fact, that of ‘autonomous technology’ is a dangerous misnomer; misnomer because it tends to obfuscate the fact that so-called autonomous systems work only thanks to the support and constraints offered by human and social interaction, first of all human knowledge, data, and social infrastructure like language, rules, practices (Stilgoe, 2017); dangerous because it promotes the false image of a technological progress that ‘doesn’t need us’ (Joy, 2000) thereby justifying the shift of capital from its source – human mental and social capacities – to other places, like the financial market.⁷

Consider the so-called sharing economy: Uber, Airbnb etc. The idea behind it is not new, and very similar initiatives have existed for a relatively long time. What makes today’s versions so successful is, in addition to the widespread availability of smartphones and other connected devices, the effectiveness of rating systems in creating trust among strangers (Calo & Rosenblat, 2017). Once again, one crucial element in the production of digital wealth is the very human capacity to assess and describe one’s individual experience and the willingness to share it with others, and the very human capacity to interpret, assess and trust the others’ assessment.

Not only should human mental and social capabilities be protected because of their intrinsic value as a constitutive element of human freedom and well-being (the capability approach point); they should also be rewarded, incentivised, and economically supported as one major source of economic value. In other words, human capabilities should also be explicitly recognised and rewarded in their role of pushing social, economic and technological progress.

Application #3: how can mental capital be rewarded?

How to do so? In what follows, we outline some potential options, yet to be assessed and developed in more details. These are presented, as it were, in order of proximity from the existing world, from the closer to the farthest, that is from the least to the most radical. As it happens, the most radical proposal may turn out to be, all things considered, the most appropriate as well as the most feasible one.

Rewarding audience labour

A first way to reward mental capital would be that of incentivising or sponsoring the creation of new digital platforms in which part of the capital freed by digital-mediated productivity flow back in the pockets of a broader range of persons who have participated in the sharing of information, knowledge, experience in online platforms, for instance, users and reviewers of digital platforms. This has been discussed under the label of ‘audience labour theory’ (Nixon, 2015; Smythe, 1977). In this context, Fischer (2015) presents and

analyses a class-action lawsuit of Facebook users against the company's use of their contents as sponsors for products in ads, which appear on their friends' News Feed. Fischer reports that users involved in the lawsuit explicitly attempt to redefine their participation in social media in terms of work: they demand more control over the information they generate and ownership over a greater share of the surplus value they produce.⁸

This approach has the merit of explicitly recognising and trying to reward the importance of the cognitive and social capabilities of the persons contributing to the success of digital online platforms, but it also presents some shortcomings, especially if applied to the so-called 'sharing economy'. First, this system would benefit only those who already possess specific kinds of skills or capabilities, for instance, the capacity and motivation to write and post helpful reviews on the digital platforms. Second, the current, largely unregulated, social interaction on digital platforms such as the reputation systems on apps has been recognised as exacerbating existing biases and therefore, in its current versions, it may also end up discriminating or marginalising particular groups of workers (Schor & Attwood-Charles, 2017). These problems might possibly be solved. However, from a broader perspective, incentivising the 'sharing economy' in its current capitalistic and under-regulated form may also exacerbate the existing differences in income as well as in the possession of skills and capabilities. Early results seem to show that actors who offer services on digital platforms are often highly skilled and educated persons who already have a (full-time) job and want to increase their regular income. Often, they even engage in manual works for which they are overqualified, like cleaning or moving, thereby making it more difficult from low-skilled workers to get a job. This has been called the problem of 'crowding out' persons in traditional jobs to the advantage of those who have the resources and skills to thrive in the sharing economy (Schor, 2017).

Autonomy-respecting social credit systems

Not everybody can and want to be an entrepreneur or a contributor in the new sharing economy, but that does not mean that they do not want to develop and deploy their physical, intellectual and social capabilities. They may want to engage in sport, art, cultural, educational, assistive activities at the local level, or in remote. Technology offers the opportunity to better tracking persons activities and to better coordinate the distribution of benefits and rewards. However, these technical opportunities seem currently to be deployed in ways that do not promote intellectual freedom. The most (in) famous example is the Social Credit System Project in China, by which people activities are constantly and systematically tracked by the State or the

municipalities in order to feed a scoring system of individuals, which is in turn used to decide on the distribution (and withdrawal) of benefits, fines and punishments (Dai, 2018).

However, digital systems of acquisition and processing of personal data may be devised to create a different system of social credit or even currency, to be used to keep track of the activities which persons perform outside traditional employment (Helbing, 2014; Marchant et al., 2014). Admittedly, even system created to achieve desirable goals may turn out to be tools to 'profile, police, and punish the poor' (Eubanks, 2018), and it may certainly be not easy to create a (state-governed) system that is at the same time effective in tracking, promoting and rewarding intellectual and social activities while at the same time protecting persons from the potentials harms coming from massive surveillance systems.⁹

Funding intellectual and social activities

In the end, it may then well be that the most revolutionary proposal to reward mental capital in the age of Artificial Intelligence may turn out also to be the simplest and oldest one: dramatically increasing the funding for educational, intellectual, scientific, social, leisure etc. activities. This would be a proposal similar in spirit but different in practice than the universal basic income discussed above. Similar insofar as the capital freed by AI-driven productivity growth would be used to provide finance for all people to invest in their own goals, capacities, talents, and interests. Different insofar as this would not be done by only providing individual persons with a (modest) sum of money. The capital should be rather used to fund targeted individual and social activities and to create free spaces and opportunities where these activities can be incentivised and supported.

Conclusion

Whether or not it will create a net job loss (aka technological unemployment), Artificial Intelligence and digital technologies will change the nature of work, and will have a deep impact on people's work lives. New political action is needed to govern this transition. In this paper we have claimed that also new philosophical concepts are needed, if the transition has to be governed responsibly and in the interest of everybody. The paper has outlined a general normative framework to make sense of- and address the issue of human freedom and justice in the age of AI at work. The framework is based on four ideas. First, in general freedom and justice cannot be achieved by only protecting existing jobs as a goal in itself, inviting persons to find ways for to remain relevant in a new machine-driven world, or offering financial compensation to those who are (permanently) left unemployed, for instance, via a Universal Basic Income. We should rather prevent technological

unemployment and the worsening of working condition to happen, as a result of a Responsible Innovation approach to technology, where freedom and justice are built into the technical and institutional structures of the work of the future. Second, more in particular, we have argued, freedom and justice may be best promoted by a politics and an economics of technology informed by the recognition of different virtues and values as constitutive of different activities, following a Walzerian ('spheres of justice') approach to technological and institutional design, possibly integrated by a virtue ethics component. Third, we have proposed that freedom and justice require that the development of individual capabilities and human agency are promoted by design. Finally, we have argued that, as a matter of fact, the mental capabilities of persons involved in the design and use of digital technology are still a crucial component of the value-creation in the (future) digital economy. These must therefore explicitly be recognised as part of a capital to be rewarded and promoted. Some first examples of domains of application of these ideas have been robotics in healthcare, citizen science, and digital platforms.

On a theoretical level, future work will have to further develop and integrate the framework and assessing its implications for the design, and introduction of other digital technologies in other work domains. On a more practical level, future interdisciplinary work will have to translate the theoretical framework into actionable design- and policy principles for specific technological and work domains.

Notes

1. See some similar points made in (Sage & Diamond, 2017, pp. 27-29). Some have also argued that it is for this reason that Basic Income is popular also among some tycoons of the Silicon Valley: it is a measure that does not directly threaten their ideology, power and business model, but rather delegates to the State the responsibility to manage their social effects (Tarnoff, 2016).
2. The general need of a more pro-active governance of innovation has been recognised also by (Acar et al., 2019) according to whom it is important to recognise that some constraints are good for innovation and (Mazzucato, 2018) according to whom policies should be devised that address big societal challenges by co-creating and shaping markets, not only by fixing them.
3. This 'resourcist' approach is also endorsed by other pre-eminent philosophers of justice, such as G.A. Cohen (1989) or Ronald Dworkin (2002).
4. On this point Philip Pettit and neo-republicans would concur.
5. On human values and agency in relation to care robots see also e.g. Coeckelbergh (Coeckelbergh, 2010) and Borenstein and Pearson (2010, 2010).
6. It's the problem of 'adaptive preferences', which has constituted one of the driving forces for the development of Sen's capability approach (Teschl & Comim, 2005).

7. A broader point in the same spirit is made by Sam Bowles (2016), when he shows that, in general, businesses, markets and other complex organizations cannot flourish without the presence of pro-social attitudes of human actors that are not accounted for in the balance sheets: trust, goodwill, selflessness.
8. See also Dirk Helbing's (2014) proposal to create an automated micropayment system based on text mining algorithms to spot re-use and duplication of texts and ideas and rewarding the authors of digital original, creative and innovative activities, without relying on intellectual property regimes or any other individual negotiations.
9. See also the recent debates on the COVID19 tracking apps for a similar problem.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Jeroen van den Hoven  <http://orcid.org/0000-0003-2376-3185>

References

- Abbott, R., & Bogenschneider, B. (2018). Should robots pay taxes? Tax policy in the age of automation. *Harvard Law & Policy Review*, 12, 145–176. <https://doi.org/10.2139/ssrn.2932483>
- Acar, O. A., Tarakci, M., & Van Knippenberg, D. (2019). Why constraints are good for innovation. *Harvard Business Review*. <https://hbr.org/2019/11/why-constraints-are-good-for-innovation>
- Alstott, A. (2001). Good for women. In J. Cohen, J. Rogers, & P. Van Parijs (Eds.), *What's wrong with a free lunch?* Boston Beacon Press.
- Amoroso, D., & Tamburrini, G. (2019). What makes human control over weapon systems “meaningful”? *ICRAC Working Paper Series #4*.
- Anderson, E. (2017). *≈*. In *Private government: how employers rule our lives (and why we don't talk about it)*. Princeton University Press.
- Barrie-Jester, A. M. (2016). What went wrong in flint. *Fivethirtyeight*. <https://fivethirtyeight.com/features/what-went-wrong-in-flint-water-crisis-michigan/>
- Bauwens, M. (2012). Blueprint for P2P society: The partner state & ethical economy. *Shareable.net* <http://www.shareable.net/blog/blueprint-for-p2psociety-the-partner-state-ethical-economy>
- Bidananure, J. U. (2019). The political theory of universal basic income. *Annual Review of Political Science*, 22(1), 481–501. <https://doi.org/10.1146/annurev-polisci-050317-070954>
- Birnbaum, S., & De Wispelaere, J. (2020). Exit strategy or exit trap? Basic income and the ‘power to say no’ in the age of precarious employment. *Socio-Economic Review* 19 3, 909–927. <https://doi.org/10.1093/ser/mwaa002>
- Borenstein, J., and Pearson, Y. (2010). Robot caregivers: harbingers of expanded freedom for all? *Ethics Inf Technol*, 12, 277–288. <https://doi.org/10.1007/s10676-010-9236-4>

- Bowles, S. (2016). *The Moral Economy: Why Good Incentives Are No Substitute for Good Citizens*. Yale University Press.
- Brynjolfsson, E., & MacAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. Norton and Company.
- Calnitsky, D. (2017). Debating basic income. *Catalyst*, 1(3), <https://catalyst-journal.com/2017/12/debating-basic-income>
- Calo, R., & Rosenblat, A. (2017). The taking economy: Uber, information, and power. *Columbia Law Review*, 117(6), 1623–1690. <https://doi.org/10.2139/ssrn.2929643>
- Carter, I. (2008). Positive and negative liberty. *Stanford Encyclopedia of Philosophy*, 363 (9410), 1–12. [https://doi.org/10.1016/S0140-6736\(04\)15657-2](https://doi.org/10.1016/S0140-6736(04)15657-2)
- Casassas, D. (2008). Basic income and the republican ideal: Rethinking material independence in contemporary societies. *Basic Income Studies*, 2(2), 2. <https://doi.org/10.2202/1932-0183.1081>
- Coeckelbergh, M. (2010). Health Care, Capabilities, and AI Assistive Technologies. *Ethic Theory Moral Prac*, 13, 181–190. <https://doi.org/10.1007/s10677-009-9186-2>
- Cohen, G. A., (1989). On the Currency of Egalitarian Justice. *Ethics*, 99(4), 906–944.
- Dai, X. (2018). Toward a reputation state: The social credit system project of China. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3193577>
- Danaher, J. (2017). Will life be worth living in a world without work? Technological unemployment and the meaning of life. *Science and Engineering Ethics*, 23(1), 41–64. <https://doi.org/10.1007/s11948-016-9770-5>
- Dworkin, R. (2002). *Sovereign virtue: the theory and practice of equality*. Harvard University Press.
- Elgarte, J. M. (2008). Basic income and the gendered division of labour. *Basic Income Studies*, 3(3), 3. <https://doi.org/10.2202/1932-0183.1136>
- Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.
- Fisher, E (2015). Class struggles in the digital frontier: audience labour theory and social media users. *Information, Communication & Society*, 18(9), 1108–1122. <https://doi.org/10.1080/1369118X.2015.1018300>
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People —An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Ford, M. (2015). *Rise of the robots: Technology and the threat of a jobless future*. Basic Books.
- Foucault, M. (1995). *Discipline and punish: the birth of the prison* (2nd Vintag). Vintage Books.
- Frank, R. H., and Cook, P. J (1996). *The Winner-Take-All Society: Why the Few at the Top Get So Much More Than the Rest of Us*. Penguin.
- Frediani, A. A., & Boano, C. (2012). *Processes for just products: The capability space of participatory design*. Springer. https://doi.org/10.1007/978-94-007-3879-9_12
- Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254–280. <https://doi.org/10.1016/j.techfore.2016.08.019>
- Friedman M. (1962). *Capitalism and Freedom*, Chicago University Press.
- Fuchs, C., & Seignani, S. (2013). What is digital labour? What is digital work? What's their difference? And why do these questions matter for understanding social media? *TripleC*, 11(2), 237–293. <https://doi.org/10.31269/triplec.v11i2.461>

- Gheaus, A., & Herzog, L. (2016). The goods of work (other than money!) *Journal of Social Philosophy*, 47(1), 70–89. <https://doi.org/10.1111/josp.12140>
- Gourevitch, A., & Stanczyk, L. (2018). The basic income illusion. *Catalyst*, 1, 4.
- Gourevitch, A. (2013). Labor republicanism and the transformation of work - Alex Gourevitch, 2013. *Political Theory*, 41(4), 591–617. <https://doi.org/10.1177/0090591713485370>
- Guimarães Pereira, Â., & Saltelli, A. (2017). Post-normal institutional identities: Quality assurance, reflexivity and ethos of care. *Futures*, 91, 53–61. <https://doi.org/10.1016/j.futures.2016.11.009>
- Habermas, J. (1987). *The theory of communicative action. Vol. II: Lifeworld and system*. Beacon.
- Hacker, J. S., & Pierson, P. (2010). *Winner-take-all politics: How Washington made the rich richer—and turned its back on the middle class*. Simon and Schuster.
- Helbing, D. (2014). Economy 4.0 and digital society: The participatory market society is born (Chapter 8 of digital society). *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2539330>
- Joy, B. (2000, January). Why the future doesn't need us. *Wired*. <https://www.wired.com/2000/04/joy-2/>
- Jütten, T. (2011). The colonization thesis: Habermas on reification. *International Journal of Philosophical Studies*, 19(5), 701–727. <https://doi.org/10.1080/09672559.2011.629672>
- Keynes, J. M. (1930). Economic possibilities for our grandchildren. In *Essays in persuasion* (pp. 358–373). Harcourt Brace.
- King, M.L. (1967). *Where Do We Go from Here: Chaos or Community?* Harper and Row.
- Little, D. (2011). Understanding society: Sen on well-being. <https://understandingsociety.blogspot.com/2011/10/sen-on-well-being.html>
- Loi, M., Dehaye, P. O., & Hafen, E. (2020). Towards Rawlsian 'property-owning democracy' through personal data platform cooperatives. *Critical Review of International Social and Political Philosophy*, 1–19. <https://doi.org/10.1080/13698230.2020.1782046>
- Lovett, F. (2018). Republicanism. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy (Summer 201)*. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/entries/republicanism/>
- Marchant, G. E., Stevens, Y. A., & Hennessy, J. M. (2014). Technology, unemployment & policy options: Navigating the transition to a better world. *Journal of Evolution and Technology*, 24(1), 26–44. <http://jetpress.org/v24/marchant.htm>
- Mazzucato, M. (2018). Mission-oriented innovation policies: Challenges and opportunities. *Industrial and Corporate Change*, 27(5), 803–815. <https://doi.org/10.1093/icc/dty034>
- Mirowski, P. (2017). Against citizen science. *Aeon*. <https://aeon.co/essays/is-grassroots-citizen-science-a-front-for-big-business>
- Naastepad, C. W. M., & Mulder, J. M. (2018). Robots and us: Towards an economics of the 'Good Life'. *Review of Social Economy*, 76(3), 302–334. <https://doi.org/10.1080/00346764.2018.1432884>
- Nixon B. (2015). The Exploitation of Audience Labour: A Missing Perspective on Communication and Capital in the Digital Era. In: Fisher E., and Fuchs C. (eds). *Reconsidering Value and Labour in the Digital Age*. Dynamics of Virtual Work Series. London: Palgrave Macmillan. https://doi.org/10.1057/9781137478573_6
- Noble, D. F. (2011). *Forces of production: A social history of industrial automation*. Routledge.

- Nussbaum, M. C. (2011). *Creating capabilities*. Harvard University Press.
- Oosterlaken, I. (2013). *Taking a capability approach to technology and its design: A philosophical exploration*. Delft University of Technology.
- Paine, T. 1797. *Agrarian Justice*. In Paine, T. (1955). *Collected Writings*. Literary Classics of the United States.
- Pellow, D. N., & Fischer, F. (2002). Citizens, experts, and the environment: The politics of local knowledge. *Contemporary Sociology*, 31(5), 578. <https://doi.org/10.2307/3090060>
- Pettit, P. (1996). Freedom as Antipower. *Ethics*, 106(3), 576–604. <https://doi.org/10.1086/233648>
- Pettit, P. (2008). A republican right to basic income? *Basic Income Studies*, 2(2), 2. <https://doi.org/10.2202/1932-0183.1082>
- Pettit, P. (2012). Legitimacy and justice in republican perspective. *Current Legal Problems*, 65(1), 59–82. <https://doi.org/10.1093/clp/cus016>
- Pogge, T. W. (2008). *World poverty and human rights*. Polity Press.
- Qizilbash, M. (2016). Some reflections on capability and republican freedom. *Journal of Human Development and Capabilities*, 17(1), 22–34. <https://doi.org/10.1080/19452829.2015.1127217>
- Rawls, J. (1971). *A theory of justice*. Harvard University Press.
- Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D., & Jacobs, D. (2012). Dynamic changes in motivation in collaborative citizen-science projects. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, 217–226. <https://doi.org/10.1145/2145204.2145238>
- Sage, D., & Diamond, P. (2017). *Europe's new social reality: The case against universal basic income*. <https://policynetwork.org/wp-content/uploads/2017/09/Europes-new-social-reality-the-case-against-UBI.pdf>
- Santoni de Sio, F., & van Wynsberghe, A. (2016). When should we use care robots? The nature-of-activities approach. *Science and Engineering Ethics*, 22(6), 1745–1760. <https://doi.org/10.1007/s11948-015-9715-4>
- Scharff, R. C., & Dusek, V. (2014). *Philosophy of technology: The technological condition: An anthology*. Wiley.
- Schor, J. B., & Attwood-Charles, W. (2017). The “sharing” economy: Labor, inequality, and social connection on for-profit platforms. *Sociology Compass*, 11(8), e12493. <https://doi.org/10.1111/soc4.12493>
- Schor, J. B. (2017). Does the sharing economy increase inequality within the eighty percent?: Findings from a qualitative study of platform providers. *Cambridge Journal of Regions, Economy and Society*, 10(2), 263–279. <https://doi.org/10.1093/cjres/rsw047>
- Sen, A., & Williams, B. (1982). Introduction: Utilitarianism and beyond. In *Utilitarianism and beyond* (pp. 1–22). Cambridge University Press. <https://doi.org/10.1017/CBO9780511611964.002>
- Sen, A. (2009). *The idea of justice*. Allen Lane & Harvard University Press.
- Sen, A. (2018). *Collective choice and social welfare*. Harvard University Press.
- Sharkey, A., & Sharkey, N. (2012). Granny and the robots: Ethical issues in robot care for the elderly. *Ethics and Information Technology*, 14(1), 27–40. <https://doi.org/10.1007/s10676-010-9234-6>
- Smythe, D.W. (1977). Communications: Blindspot of Western Marxism. *Canadian Journal of Political and Social Theory*, 1(3), 1–27.

- Stewart, A., & Stanford, J. (2017). Regulating work in the gig economy: What are the options? *The Economic and Labour Relations Review*, 28(3), 420–437. <https://doi.org/10.1177/1035304617722461>
- Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9), 1568–1580. <https://doi.org/10.1016/j.respol.2013.05.008>
- Stilgoe, J. (2017). Machine learning, social learning and the governance of self-driving cars. *Social Studies of Science*, 48 (1) , 25–56. <https://doi.org/10.1177/0306312717741687>
- Tarnoff, B. (2016). Tech billionaires got rich off us. Now they want to feed us the crumbs. *The Guardian*. <https://www.theguardian.com/technology/2016/may/16/universal-basic-income-equality-tech-silicon-valley>
- Teschl, M., & Comim, F. (2005). Adaptive preferences and capabilities: Some preliminary conceptual explorations. *Review of Social Economy*, 63(2), 229–247. <https://doi.org/10.1080/00346760500130374>
- Thompson, M. J. (2013). Reconstructing republican freedom: A critique of the neo-republican concept of freedom as non-domination. *Philosophy & Social Criticism*, 39 (3), 277–298. <https://doi.org/10.1177/0191453712473081>
- Valentini, L. (2012). Ideal vs. Non-ideal theory: A conceptual map. *Philosophy Compass*, 7(9), 654–664. <https://doi.org/10.1111/j.1747-9991.2012.00500.x>
- Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*. Oxford University Press.
- Van Den Hoven, J., Vermaas, P. E., & Van de Poel, I. (2015). . In *Handbook of ethics, values, and technological design: Sources, theory, values and application domains*. Springer. <https://doi.org/10.1007/978-94-007-6970-0>
- van Den Hoven, J. (2013). Value sensitive design and responsible innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible innovation* (pp. 75–83). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118551424.ch4>
- Van der Veen, R. (1998). Real freedom versus reciprocity: Competing views on the justice of unconditional basic income. *Political Studies*, 40(6), 140–163. <https://doi.org/10.1111/1467-9248.00134>
- Van Donselaar, G. (2009). *The right to exploit: Parasitism, scarcity, basic income*. Oxford University Press.
- van Parijs, P. (Ed.). (1992). *Arguing for basic income: Ethical foundation for a radical reform*. Verso Books.
- van Parijs, P., & Vanderborght, Y. (2019). *Basic income: A radical proposal for a free society and a sane economy*. Harvard University Press.
- van Parijs, P. (1995). In *Real freedom for all: what (if anything) can justify capitalism?* Clarendon Press.
- Van Wynsberghe, A. (2015). *Healthcare robots: Ethics, design and implementation*. Routledge.
- Walker, M. (2014). BIG and technological unemployment: Chicken little versus the economists. *Journal of Evolution and Technology*, 24(1), 5–25. 15410099. <https://jetpress.org/v24/walker.htm>
- Walzer, M. (1983a). *Spheres of justice: A defense of pluralism and equality (Nachdr.)*. Basic Books.
- Walzer, M. (1983b). *Spheres of justice: A defense of pluralism and equality*. Basic Books.
- Walzer, M. (1984). Liberalism and the art of separation. *Political Theory*, 12(3), 315–330. <https://doi.org/10.1177/0090591784012003001>

- Warren, D. M. (1991). Using indigenous knowledge in agricultural development. *World Bank - Discussion Papers*, 127.
- Wildschut, D. (2017). The need for citizen science in the transition to a sustainable peer-to-peer-society. *Futures*, 91, 46–52. <https://doi.org/10.1016/j.futures.2016.11.010>
- Willcocks, L. (2020). Robo-Apocalypse cancelled? Reframing the automation and future of work debate. *Journal of Information Technology*, 35(4), 286–302. <https://doi.org/10.1177/0268396220925830>
- Winner, L. (1980). Do artifacts have politics? *Daedalus*, 109(1), 121–136. <https://doi.org/10.1016/j.futures.2016.11.010>
- Wright, E. O. (2006). Basic income as a socialist project. *Basic Income Studies*, 1(1), 1. <https://doi.org/10.2202/1932-0183.1008>
- Young, I. M. (1990). *Justice and the politics of difference*. Princeton University Press.
- Zwolinski, M. (2019). A hayekian case for free markets and a basic income. In M. Cholbi, and M. Weber (Eds.), *The future of work, technology, and basic income*. Routledge. 7–26.