



Critical Review of International Social and Political Philosophy

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/fcri20>

Towards Rawlsian 'property-owning democracy' through personal data platform cooperatives

Michele Loi , Paul-Olivier Dehaye & Ernst Hafen

To cite this article: Michele Loi , Paul-Olivier Dehaye & Ernst Hafen (2020): Towards Rawlsian 'property-owning democracy' through personal data platform cooperatives, Critical Review of International Social and Political Philosophy, DOI: [10.1080/13698230.2020.1782046](https://doi.org/10.1080/13698230.2020.1782046)

To link to this article: <https://doi.org/10.1080/13698230.2020.1782046>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 16 Jun 2020.



Submit your article to this journal [↗](#)



Article views: 1346



View related articles [↗](#)



View Crossmark data [↗](#)

Towards Rawlsian ‘property-owning democracy’ through personal data platform cooperatives

Michele Loi^a, Paul-Olivier Dehaye^b and Ernst Hafen^c

^aDigital Society Initiative, University of Zurich, Zurich, Switzerland; ^bPersonalData.IO, Geneva, Switzerland; ^cInstitute for Molecular Systems Biology, ETH Zurich, Zurich, Zurich Switzerland

ABSTRACT


This paper supports the personal data platform cooperative as a means of bringing about John Rawls’s favoured institutional realisation of a just society, the property-owning democracy. It describes personal data platform cooperatives and applies Rawls’s political philosophy to analyse the institutional forms of a just society in relation to the economic power deriving from aggregating personal data. It argues that a society involving a significant number of personal data platform cooperatives will be more suitable to realising Rawls’s principle of fair equality of opportunity.

KEYWORDS Big data; justice; cooperatives; equality of opportunity; John Rawls

Introduction

Personal data is sometimes described as the new oil of the world economy. If it is true that personal data, especially in the context of so-called big data, plays the role of an important economic resource affecting the futures of individuals and countries, then the normative questions concerning big data are much broader than privacy and data protection. In fact, since they concern the distribution and creation of resources, they are questions of distributive justice. When evaluating the functioning of markets, most economists consider only the value of efficiency. In contrast, this paper applies John Rawls’s political philosophy, which proposes a concept of socio-economic justice that goes beyond efficiency. We advocate the personal data platform cooperative (PDPC) as a means to bring about Rawls’s favoured institutional realisation of justice, the property-owning democracy (POD).

We argue that a certain form of inequality in the data economy is a violation of Rawls’s principle of fair equality of opportunity (FEO). As Rawls argues, FEO is violated by economic institutions that place productive assets under the control of a relatively small sector of society. Rawls’s idea of the POD is intended to address this problem. Here, we explore a parallel

CONTACT Michele Loi  michele.loi@uzh.ch

© 2020 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 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

problem in the data economy¹ and argue that an institutional order analogous to a POD for data can be achieved with PDPCs.

The issue can also be described in terms of the pragmatic egalitarian political principle that productive assets should be distributed as widely as possible, as opposed to concentrated in few hands. The idea of predistribution has been discussed in both academic and policy circles as an alternative to ‘tax and redistribute’ welfare state capitalism (Hacker, 2015; Kerr, 2015; O’Neil, 2017). Predistribution aims ‘to focus on market reforms that encourage a more equal distribution of economic power and rewards even before government collects taxes or pays out benefits’ (Hacker, 2015). Predistributive policy is supported by the pragmatic argument that excessive reliance on tax-and-spend redistribution ‘fosters backlash, making taxes more salient and feeding into the conservative critique that government simply meddles with “natural” market rewards’ (Hacker, 2015). Predistribution models emphasise the need to establish market correctives that achieve egalitarian outcomes by distributing control over productive resources, considering ex-post correctives to be a second-best alternative.²

This paper is not meant only as an addition to Rawlsian scholarship. Its conceptual innovations, especially the idea of pre-distributing the productive assets of the data economy, may appeal to those who are not committed to a Rawlsian view of social justice. Thus, this paper contributes to the broader transdisciplinary debate on the politics of big data.³

In this discussion, we consider only personal data, which we define customarily as any data related to an identified or identifiable individual. It is important to stress that, as anticipated in EU law,⁴ the definition of personal data will evolve dynamically because data’s identifiability depends on what other data is available and the evolution of technology.

Moreover, our purpose here is not to provide a blueprint for the PDPC that is ready to be implemented. Since we aim to situate this proposal in relation to the politics of data, we focus on the concept of the PDPC and what distinguishes it from other models rather than on the details of its operations.

Finally, our proposal differs from others that seek to reorganise the data economy to promote justice (Cheneval, 2018) in that, p. 1) it does not promote the full ownership and control of individual data and 2) it focuses on the unequal distribution of what Rawls calls ‘prerogatives of authority and responsibility’ in the data economy rather than the unequal income deriving from them. The proposal differs from those that stress the personal ownership of data in that it rests on the idea of the collective ownership and control of data-collection infrastructure. Moreover, our proposal frames collective ownership as a way to distribute the power to shape the online environments that subtly encourage (or ‘nudge’) citizens to share their data for further uses.

The paper unfolds as follows: [Section 1](#) presents Rawls’s theory of justice and the POD as its institutional realisation in the context of the Rawlsian

criticism of welfare state capitalism. It also extends this criticism to the current data economy. [Section 2](#) presents our proposal for organising the data economy, the personal data platform cooperative (PDPC). [Section 3](#) describes a society in which PDPCs are significant economic actors that offer opportunities to control large data assets to anyone who chooses so and with a stake in that particular data. It argues that such a society would approximate a form of POD and would be more likely to realise an approximation of FEO, at least in the data domain. [Section 4](#) explains how the PDPC proposal represents an innovation relative to the privacy self-management paradigm.

Property-owning democracy and Rawlsian principles of justice

Today, the internet is dominated by large companies that produce and control vast quantities of personal data as a collateral effect of providing their services through the internet. The primary technological innovations responsible for this state of affairs are the internet and the smartphone. These technologies provide constant, global, real-time access to a wide variety of services such as maps, blogs, videos and internet searches. As a side effect (from the user's perspective)⁵ of the interactions between platform companies and their customers, formidable amounts of data are collected.

The data produced and controlled by platform owners are considered a 'new asset class' (World Economic Forum & Bain & Company, Inc, 2011). The ability to control and benefit from such assets is marked by significant inequalities. First, there is an inequality in the ability to collect and control these personal data assets, as the internet contains only a few gatekeepers (Google, Facebook, etc.) who are uniquely positioned to do so. The gatekeepers' unique position derives from a combination of various network effects that make it difficult, if not impossible, to compete against the first company that significantly benefits from them. For example, it is difficult for a company to compete against Facebook by offering a similar product if it starts with a far smaller user base. In the case of social networks, the number of users determines the number of potential 'friends' each user can reach and thus adds value to the service.⁶ Similarly, Google Search benefits from a host of interlocking network effects: marketplace network effects (advertisers affiliated with Google can access individuals with the best profiles, while each advertiser contributes to profiling), data network effects (more data makes it easier to build an ecosystem of services around each user), recruiting network effects (more data supports better services that attract more people), and feedback network effects (the behaviour of users tells Google which search results are selected after typing a given search key, thus helping Google to identify the most fitting search results).⁷ Due to these network effects, many internet services markets (e.g., the search or social network markets) tend to be winner-takes-all.

In today's data exchange market, free services are exchanged for personal data. Typically, users are not fully aware of the various ways in which their data are used and cannot conceptualise the real economic value of the assets they produce. Users are only offered intangible and unquantified transactions with their data, which systematically and significantly complicates the rise of more favourable economic arrangements (Haynes & Carolyn Nguyen, 2014). Even if an individual were aware of the economic potential of their data, their control over the data would be too limited to make it economically rational⁸ for a single individual to bargain with data collectors.

Let us grant for the sake of argument that the current data economy is as efficient as possible. That is, let us suppose that users of digital services obtain at least as much utility from free services as they could obtain in any arrangement requiring users to pay for services and companies to pay for users' data. From the point of view of efficiency, there can be no argument against such an arrangement. However, Rawls's argument for a POD reveals an important flaw in the current model: the concentration of economic power and the resulting concentration of political influence.

For a long time, this aspect of Rawls's political theory attracted very little attention by commentators. The discussion has reflected a kind of standard assumption, more often implied than stated, that a just arrangement could be achieved by an ordinary, North European-style welfare-state capitalist society. The standard assumption was that Rawls's society was another vision of a 'slightly imaginary Sweden,' to use Robert L. Heilbroner's famous expression.

The last decade has witnessed a resurrection of Rawls's critical stance against welfare state capitalism. Rawls admits that the least-advantaged individuals in this social configuration are better off compared to laissez-faire capitalism but claims that the system fails to satisfy all the requirements of his theory of justice. In societies where the means of production are privately owned, Rawls affirms, justice can only be achieved by a POD. While capitalism as we know it gives all citizens sufficient opportunities to achieve a decent life, it also enables a small class of property owners 'to have a near monopoly of the means of production,' even when capitalism is complemented by the generous welfare system of an 'imaginary Sweden' (Rawls, 2001, p. 139). Thus, welfare state capitalism might lead to self-perpetuating economic and political inequalities (O'Neill, 2009). The positive features of welfare state capitalism include public support for various forms of social insurance, education and health care, which ensure that the basic needs of all citizens are adequately met. However, in welfare state capitalism, significant inequalities between citizens born in different strata of the population may persist and even grow indefinitely. Rawls doubts that capitalism can benefit the least-advantaged members of society through mere trickle-down effects. He also doubts that welfare state capitalism can benefit the least-advantaged to the degree required by the principles of his theory of justice. There are two main

reasons, in Rawls's view, why welfare state capitalism fails to provide the institutional bases of a just society:

- (a) Citizens do not have similar opportunities to access positions that are characterised by the same prerogatives of authority and responsibility. This is a violation of the Rawlsian principle of FEO.⁹ When some have access to initial productive resources (e.g., through gifts and inheritance) that others lack, it is extremely difficult to level the playing field (through the redistributive strategies of welfare state capitalism) in the competition for social positions.
- (b) Redistributive policies, which may often be instrumental to satisfying the difference principle,¹⁰ cannot be implemented when they conflict with the interests of powerful economic elites (O'Neill, 2009).¹¹

The FEO principle is extremely demanding. It would be satisfied in the current economy only if the state could neutralise the influence of social positions on individuals' prospects of obtaining positions characterised by significant power and control of economic assets, including in the data assets. Rawls is keenly aware that the political target specified by FEO is extremely difficult to realise under ordinary capitalist institutions, including those that effectively protect and promote the interests of the economically worst-off members of society through state-supported social welfare. The reason that FEO is so hard to realise in a highly unequal society is that the initial advantages of those born in socially and economically advantaged social classes cannot be adequately compensated by welfare services. Even if the competition for scarce educational resources can be isolated from the influence of unequal material circumstances, the inequality-generating effects of intergenerationally transmitted cultural and social capital are much harder to mitigate (Bowles, 2005; Brighouse & Swift, 2006; Savage et al., 2015).

Rawls's contrasting proposal is that of a property-owning democracy (POD). An economic arrangement is a POD if the economic system as a whole

tries to disperse the ownership of wealth and capital, and thus to prevent a small part of society from controlling the economy and thus indirectly political life itself. [...] The idea is not simply to assist those who lose out through accident or misfortune (although this must be done), but instead to put all citizens in a position to manage their own affairs and to take part in social cooperation on a footing of mutual respect under appropriately equal conditions. (Rawls, 1999, xiv–xv)

An example Rawls often gives of a social institution appropriate for a POD is the taxation of large inheritances. If taxation applies to inheritance, and capital is redistributed at the beginning of an individual's life, the economic

gap that will have to be filled by other kinds of social policy (e.g., public education) will be smaller.

Rawls argues that the principles of justice are most likely to be realised in a POD, rather than in welfare state capitalism (no matter how generous the level of welfare services). In a POD, citizens have private ownership over productive resources, but inheritance taxation prevents the concentration and accumulation of great fortunes in the hands of the few. Hence Rawls's characterisation of a POD in his penultimate work, *Justice as Fairness: A Restatement*, as a social configuration in which institutions

put in the hands of citizens generally, and not only a few, sufficient productive means for them to be fully cooperative members of a society on a footing of equality. (Rawls, 2001, p. 140)

We summarise the FEO-based argument for POD as follows: it is more difficult to reduce inequalities of opportunity between similarly talented and motivated individuals (e.g., by improving education) in societies with very large initial inequalities.

In the current data economy, very few people gain opportunities to control significant amounts of personal data. In other words, the central capability of control over personal data assets is distributed very unequally across social positions. Most users of data-based services are only able to provide consent for specific uses of their own personal data, but they have no power to determine the uses for which their consent may be asked. By contrast, a few managers and owners of large companies (e.g., Mark Zuckerberg, who still owns the majority share of his public company; Hiltzik, 2019) exert a disproportionate amount of power and control over the terms and conditions of economic exchanges involving personal data. Moreover, policies that may be in place to promote FEO (e.g., education policy) cannot close the initial gap¹² in access to positions that command substantial productive resources in the data economy.

Clearly, no plausible reading of FEO requires that all equally talented and motivated individuals have the same opportunities to exercise the authority to control every relevant domain of our life in which such authority is relevant. Indeed, there may be trade-offs in the degree to which opportunity can be equalised in various spheres of economic and social life. Mitigating inequality of opportunity in all competitions for all social positions *equally* may not be the most reasonable and just solution. However, in order for the FEO principle to have any normative salience, some inequalities in talent and motivation must be prioritised. Talent and motivation inequalities that reflect unequal social access to basic goods that are necessary to develop talents and motivations (such as education and health) are not permitted by FEO (Daniels, 1981). It is, however, difficult to determine once and for all what such basic goods are, as the definition of 'basic goods' is contingent and may differ between historical eras.¹³

We speculate that control over data assets is one of the basic goods in question. Access to the central socioeconomic good of data assets is a resource that shapes an economic and social actor's overall opportunities (affecting his or her chances of obtaining other prerogatives of authority and responsibility) and therefore especially relevant to justice. We contend that in both economics and politics, and possibly in other competitive spheres of social life, access to large data assets can be a crucial competitive advantage. Recent events such as the Cambridge Analytica scandal (Grassegger & Krogerus, 2016) show that better access to and control over large personal data assets produce competitive advantages in both economic and political competition.¹⁴

In summary, FEO is violated if individuals do not have similar core capabilities of data access and usage, which, in the current economic configuration, affect individuals' chances of success in other competitive spheres.

We thus argue, by analogy with Rawls's argument for POD, that FEO is more easily achieved if there are institutions that pre-distribute control of significant data assets more broadly. Clearly, the amount of authority available to every individual in society can never be equal to that of a CEO of an existing big data corporation. However, authority could be more widely distributed and thus diluted. In this way, inequalities of opportunity related to this central capability could be more easily mitigated, as we shall argue in [Section 3](#).

The idea of personal data platform cooperatives

How can inequalities of opportunity in the control of significant data assets be mitigated? This is an especially challenging question given that personal data are not collected by a centralised body that simply collects the data that users provide. Rather, most valuable personal data are collected by particular businesses and organisations that track users' interactions within their platforms.

The main strategy that we suggest as a means to deal with this problem begins with the observation that personal data are not a rival good.¹⁵ Digital data can be copied perfectly and inexpensively. They can be stored in different places and reused a potentially infinite number of times by anyone with the right hardware and technical and legally authorised access to the original information (or a copy of it). Furthermore, data collected by different businesses, organisations and platforms can often be linked to the same individual (again, with appropriate legal permission). Thus, data can be aggregated in a way that enhances their social and economic usefulness and productivity (as will

be discussed in [Section 3](#) [Personal data platform cooperatives and property-owning democracy]).

At the highest level of abstraction, personal data platform cooperatives (PDPC) provide a governance model for personal data aggregators and providers of data-driven services based on two essential features:

- (A) a personal data management platform (PDMP) empowering individuals to collect, aggregate and control (copies of) their personal data from different sources (e.g., genomic data, e-health records, and e-commerce data), enabling clients to choose what data to share and with whom; and
- (B) democratic procedures that enable cooperative members to make collective choices concerning:
 - (a) general data analytics capabilities, policies and ethical codes for data transactions and services delivered through the PDMP; and
 - (b) the deployment of surplus deriving from the secondary utilisation of data by third parties (e.g., for research, industry, or marketing purposes), once all costs associated with running the platform are deduced.

Cooperative members collectively exercise power over what kinds of data analytics capabilities, policies and ethical codes apply to data collected and controlled through the collectively owned platform. The following are examples of decisions that PDPC members would make collectively: whether genetic data should be collected by the platform, whether such data should be made accessible for commercial services in general, and whether transactions involving genetic data with insurance companies should be allowed, and if so, in what form and with what constraints. While the cooperative as a whole defines the structure and possibilities enabled by the data-sharing environment, individuals make their own decisions about whether to share their personal data.

PDPCs are designed to empower individuals to control their own personal data while limiting such power through general policies and ethical codes voted on by democratic majorities (members' general assemblies). As account holders, members control access to their data (as in a bank account, where individual data is individual money). As cooperative members, they control the sharing environment (how the data is measured, visualised, classified, and controlled) and the data-sharing limits (e.g., what type of data can be shared with specific categories of stakeholders) applicable to all members.

However, if data are like oil, PDMPs without personal data are like engines without fuel. For PDMPs to be of any use, end users must gain control over the personal data assets that they have donated, sometimes inadvertently, as

a collateral effect of previous online interactions. Here is where the right to obtain copies of one's personal data enters the picture. Such rights have been conferred, for example, by the recent EU General Data Protection Regulation (Regulation on the protection, 2016), and in particular Article 20 concerning the right to data portability (Right to data portability, 2016).¹⁶ The right to data portability is the right to obtain data collected by private service providers in machine-readable form. It obliges private entities to make copies of personal data available to the data subject herself or to another company following the subject's request. Currently, large platform corporations (e.g., Google and Facebook) comply with EU law and make it possible for their customers, including those outside the EU, to download a machine-readable copy of the personal data that they have collected.¹⁷

The idea of the PDPC departs from that of privacy self-management, which is based on notions of notice, access and consent regarding the collection, use and disclosure of personal data. The PDPC does not rely on the idea that 'people can decide for themselves how to weigh the costs and benefits of the collection, use, or disclosure of their information' (Solove, 2012, p. 1880) in isolation from a community of some sort. Still, it retains certain features of the privacy self-management concept: each member (*qua* platform user) is asked for individual consent for specific uses of his or her data. In a PDPC, personal data can never be used and shared without the explicit informed consent of the individual. The consent given to a PDPC should not be wide consent authorizing a PDPC to use all data about an individual, but consent involving more granular sharing options, which can be dynamically updated and revoked (Budin-Ljøsne et al., 2017). With these granular consent options, members can make a specific subset of data available for certain purposes but not others.

Most importantly, however, PDPCs revise the privacy self-management concept in the direction of a community-management concept of the digital environment. Each member (*qua* cooperative member) has an equal say in defining the digital experience of data collection (e.g., the nudges built into the graphic design of a platform in the user journey and the choice of available options) affecting how their data are collected and how potential future uses are presented.¹⁸ Moreover, each member has an equal say about the constraints – most importantly, the ethical constraints – on new services and sharing opportunities offered through the platform.

This can be achieved in the current legal framework by technological platforms owned by cooperatives of data subjects. Cooperative governance is based on the customary cooperative rule of one share and one vote per member. While majoritarian decisions coerce minorities into accepting a specific way of shaping the digital environment for sharing as well as the available sharing options, this form of coercion can be justified because of its limits. First of all, PDPCs are chosen and voluntarily entered into (in Rawlsian

terms, they are free associations). Second, the constraints only apply to the data collected through the platform and the transactions implemented through them. Third, each cooperative member can exercise the right to obtain a copy of the data collected through the platform in a machine-readable format and leave the cooperative. Low-cost exit rights guarantee that there is no user lock-in and protect individual autonomy from the will of majorities that happen to form within the platform. Most importantly, data portability should enable a social world populated by a *plurality* of different PDPCs, each governing large data assets, not a single PDPC acting as a global monopolist for the entire world. Individuals shall be free to leave a cooperative (demanding the cancellation of their data) and may be able to join different PDPCs at the same time.

Note that it is not essential to our argument that *all* personal data about the same individual should be controllable by an IT infrastructure owned by a PDPC. The proposal is feasible as long as there are sufficiently large datasets of unmistakably personal data, people are aware of these datasets, and there are laws allowing people to obtain a copy of their personal data. The goal is not to control all the data a person may produce about herself. It is, rather, to provide shared significant control to a community (typically of individuals sharing similar interests and values) over a significant portion of data.

Personal data platform cooperatives and property-owning democracy

In this section, we advocate PDPCs from a Rawlsian perspective by showing that a PDPC-driven data democracy (PDDD) is preferable to the current data economy, from the standpoint of FEO. The PDDD is an ideal form of the data economy in which a large proportion of the users of data-based services are members of one or another data cooperative.

The argument for PDDD is structurally analogous to Rawls's argument in support of POD. Rawls advocates a 'regime in which land and capital are widely though not presumably equally held' and '[s]ociety is not so divided that one fairly small sector controls the preponderance of productive resources' (Rawls, 1999, p. 247). A society organised in this way is more likely to satisfy FEO than welfare state capitalism, as in the latter system assets may become highly concentrated over time. By analogy, the argument for a PDDD is based on the claim that an economic regime where most data transactions are controlled by PDPCs is more likely than the current data economy to satisfy FEO, as it would broaden access to large data assets and thus mitigate inequality of opportunity to control these assets.

For the sake of simplicity, let us consider a 'pure PDDD,' defined as a society where *all* personal-data-based services are provided by PDPCs and each user of a data-based service is also a member of the PDPC that offers it.

This is, of course, not a realistic scenario. Information is a non-rival good, so cooperative users' control of one copy of their personal data through the PDPC does not eliminate the existence or continued use of other copies of the same data by other data controllers, including private companies. Moreover, PDPCs do not compete against private corporations that provide data-driven services for scarce informational resources, and they may even develop a symbiotic relationship with them. For example, users of PDPC services may transfer data to and from PDPCs and companies via data portability rights. In a mixed economy involving both private companies and PDPCs, the latter may not be predominant. Still, it is worth addressing the pure PDDD as an ideal type to see where the discussion leads. From the analysis of the pure PDDD, it may be possible to derive some indications of how a 'mixed PDDD' would work. A mixed PDDD would be a society containing many PDPCs, most of which play significant economic roles, that compete and collaborate with large private companies that of roughly equal economic importance.

How do PDPCs contribute to realising the conditions for POD in the information society? As argued in [Section 1](#) [Property-owning democracy and Rawlsian principles of justice], mere users are currently excluded from the layer of economic interactions in which profits are made: the secondary use of data. Moreover, mere users can decide whether to give their consent to terms and conditions and digital environments determining how data are collected and used; terms and conditions and digital environments are nowadays defined by companies, constrained by national and international laws, which are themselves shaped by democratic processes.

As argued in [Section 2](#) [The idea of personal data platform cooperatives], PDPC members enjoy equal votes in the general assembly. In contrast to the current data economy, where authority and control over large data assets are concentrated in apical management roles and ownership roles, a PDDD offers opportunities for social roles with an intermediate amount of control over significant aggregations of data. Namely, in their role as a cooperative member, a person would have an intermediate capability to control significant amounts of data; their degree of control would be intermediate between the capability of most individual citizens today and that of owners and managers of large data-driven companies. Unlike the owner or manager of a large data company, a PDPC member would have the amount of power and influence afforded by their single vote and their ability to persuade other cooperative members. As a result, in a pure PDDD, large portions of the population, as opposed to a small sector of society, would exert some control over the new productive asset class. A large 'middle class' of data control could thus be established.

Realistically, a mixed PDDD may still include large corporations and, within them, positions attached to greater power over aggregate data. However, this is not a fatal flaw of the model. By analogy, a POD may still include very rich

individuals (e.g., highly successful innovators and artists), but this is compatible with FEO because it does not entail that a minority of citizens control most productive resources.

Having explored the structural analogies between a PDDD and a POD, let us now turn to the argument that FEO is more likely to be realised by a PDDD than by the current data economy.

The current data economy is a society in which authority over data is distributed very unequally, and it is thus defective from the standpoint of FEO. There are few positions of authority over aggregate data, and opportunities characterised by access to large data resources are very unequally distributed. Realistically, policies that 'level the playing field' (e.g., through education), will not neutralise all inequalities of social class and birth. While opportunity inequality in the data economy could be ignored in the past, it has become more problematic as the control and exploitation of large data assets is increasingly central to success across multiple social domains. In other words, the control and exploitation of large data resources have become important determinants of other opportunities (e.g., social, economic, and political opportunities). By contrast, in a pure PDDD, authority is widely distributed and all interested persons, irrespective of their ownership of capital assets, exercise some share of authority over large aggregates of data if they both have a desire to do so and are an interested party in those data (i.e. the data in question are *their* personal data, that is, data which are about *them*). The minimal cultural means necessary to use the opportunities offered by PDPC members could be delivered by public education. This would not pose an undue burden for the state, since it is reasonable that public school should finally educate persons about the nature and societal implications of data sharing, control and aggregation. Such education is required for citizens to use digital services in an informed manner.¹⁹ While people with vastly divergent interests will not be equally drawn to participate in PDPCs, this does not count as a violation of FEO, which requires similar opportunities only among people whose motivations (and initial talents) are similar.

While our leading normative justification for the data cooperative is equality of opportunity, not efficiency, other scholars have emphasised the potential of data cooperatives to create value. Consider the network effects (discussed in [Section 1](#)) that make it difficult for similar companies to compete against Google and Facebook. Network effects are only one aspect of a broader phenomenon: the economic value of systematically aggregated data is much larger than the sum of its parts. This is true more generally, as data linkage enables 'super-additive insights' (OECD, 2014, p. 29). In other words, the ability to better contextualise data enhances the amount and quality of insight that can be derived from it. This phenomenon is the basis of the economic reason for distinguishing individual control of personal data

(which some individuals already have by virtue of data protection law) from the collective control of large data assets. PDPCs may enable the nonlinear, increasing returns of scale for the monetary, economic, and social value of data, which derives from their broad and systematic integration. As the number of records with which a single record can be compared increases, so does the value of each piece of information (OECD, 2013).

If this is true, data cooperatives could unlock the potential of personal data, which is currently unrealised because most data is collected in silos. In a PDPC, the cooperative member would provide the moral and legal authority for the integration of all his or her personal data, enabling value creation through aggregation. Moreover, it has been argued that establishing data cooperatives in the health domain could accelerate research and its clinical applications (Blasimme et al., 2018). In contrast to aggregation by external entities, individuals would retain control over their aggregated data through the data management functions of the PDMP owned and operated by the PDPC (Section 2). Finally, the unlocked value of aggregated data would enhance the power and opportunities of all cooperative members as a collective, in contrast to the concentrated power and opportunities of today's data aggregators. This would make society more just, since the latter configuration leads to great inequalities of economic opportunity and political influence, while the former arrangement mitigates them.²⁰

Beyond privacy self-management

In this section, we wish to emphasise how our model goes beyond the privacy self-management model. Since the establishment of the principles of fair information practice, privacy has been assumed to be a realisation of individual autonomy, and individual autonomy has been understood as the result of endowing individuals with more information about the possible uses of their data. However, it has become increasingly clear that this solution is not feasible given the ubiquity of data and data transactions in the contemporary internet economy (Cate, 2006). In this environment, notice and consent practices can easily lead to information overload (e.g., informed consent notifications that many users simply avoid). There are simply too many data exchanges and too many uses for individuals to be able to review everything independently; privacy notices are like a distributed denial-of-service (DDoS) attack on our brain (Hartzog, 2018). Ultimately, data use decisions are determined by the nudges that are built-in to digital platforms (Hartzog, 2018; Weinmann et al., 2016).

Online choice architectures involve nudges because the visual layout, user journey, default settings, feedback mechanisms, and user dashboards of a web platform are unavoidable design elements and often constitute nudges, intentionally or unintentionally. The attention of the user of

a digital platform is a scarce resource. When an online artefact is built, design decisions affect the default choices and the attention economy of the user (Hartzog, 2018; Weinmann et al., 2016).

Hence, we argue that a fundamental aspect of agency in the internet environment and in a big data economy is having some form of control over the choice architecture within which data transactions occur. Data cooperatives would expand and equalise opportunities to control online choice architecture. Since control over a shared choice architecture implies imposing one arrangement onto many individuals, the problem of preference aggregation arises. In a dictatorial model, a single agent, or small group of agents, e.g., a management board, imposes a choice architecture for all other agents to use. By contrast, in a democratic model, experts develop alternative options that are ultimately voted on by cooperative members. The democratic solution delivers a low level of inequality between the users and designers of data-driven services that is compatible with a desirable degree of efficiency. Mitigating inequality in the ability to shape online environments is morally important because fundamental choices concerning the default settings, consent mechanisms and visual layouts of online platforms require balancing values (e.g., a design may privilege data privacy and data security, another design may nudge people into sharing more data more broadly) whose relative importance cannot be decided by technical considerations alone (Mirsch et al., 2017; Weinmann et al., 2016). An individual's basic capability to control significant data assets also includes the higher-order capability to (democratically, collectively) shape the rules and nudges implemented in the online environments affecting some (if not all) of the data transactions in which she or he is involved.

The idea of exercising economic power through not just data choices, but also through data interfaces, realizes one peculiar facet of *predistribution in the data economy*. What gets pre-distributed, in this case, is not an ordinary economic asset (something with a clear market value), but economic power in a more abstract form: the capacity to influence the behavioral process which generates data and influences their economic value.

Conclusion

Having presented Rawls's idea of a property-owning democracy (POD) and its justification through Rawls's second principle of justice, in particular fair equality of opportunity (FEO), we have worked out the implications of this vision for the socioeconomic institutions of a society in which personal data play a significant economic role. The possibility of a POD in a technologically advanced economy is threatened by the persistent and self-reinforcing concentration of economic power by dominant parties. Today, this is expressed in the concentration of the power to shape the digital environments affecting

the collection and repurposing of data, as well as the options for allowing data use, offered to average citizens. Options and digital environments are bundled and offered as take-it-or-leave-it options.

PDPCs are institutional tools that empower citizens to have authority not only over their personal data but also, through democratic procedures, on their digital environments and the options open to them. We have argued that, in a data economy where PDPCs control a significant proportion of (personal) data-based services, FEO is more likely to be achieved. This argument is structurally analogous to Rawls's argument for a POD as the social institution most likely to satisfy FEO.

It may still be possible that a data economy with a significant role for PDPCs is not the only economic arrangement structurally analogous to a POD. Other institutions may contribute to the democratisation of power over big data. Still, the PDPC is at least one option deserving consideration.

Notes

1. Note that one may argue that the current data economy is unjust for other reasons, e.g., because it is discriminatory, opaque, monopolistic, exploitative, and manipulative (Custers et al., 2012; Pasquale, 2015). In this essay we choose to explore a particular Rawlsian, or predistributive, standpoint, as we believe it conceptualizes the existing injustice in a new light. Moreover, we do not argue that the PDPC can remove all forms of injustice in the data economy; the broader issues of injustice in the data economy deserve an analysis that would exceed the size of a journal article.
2. One important difference between the Rawlsian argument and the predistribution argument is that the former appeals to principles, while the latter is couched in pragmatic language. Predistribution theorists emphasise that, while it may be theoretically possible to achieve a synthesis of equality and efficiency through fair taxation and redistribution, such policies encounter strong resistance in practice. These theorists argue that an alternative approach – an initial redistribution of assets that otherwise contribute to generate highly unequal income over time – is at least worth exploring, as it may face less resistance.
3. The role of political philosophy in the debate has been developing while this paper was under review. See, for instance, Ferretti (2020), who considers the virtues of the platform cooperative model more generally and briefly touches on redistributing the benefits deriving from the cooperative's access to aggregate data. Our proposal falls fully within the 'organizational strategy' that Ferretti describes. In his defense of this strategy, Ferretti also draws on the work of Martin O'Neill (2009) concerning the idea of a Rawlsian POD. See also Carballa Smichowski (2016) for an alternative policy proposal to create a data commons controlled by multi-stakeholder councils.
4. Recital 26 of the GDPR holds that 'To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used [...], account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration

the available technology at the time of the processing and technological developments.’

5. In the current data economy, collecting data in order to predict and influence the (for now, mainly commercial) behaviour of internet users is for most companies the main business model, but it hides behind the appearance of being merely incidental.
6. It might be objected that Google may soon face serious competition by Bing, the Microsoft-powered search engine. This is the kind of exception that proves the rule: few companies are able to sustain the huge losses that Microsoft suffered for several years in order to get a chance to compete with Google, and even in this case, the possibility of competition only exists because Microsoft can exploit its dominance in the operating systems market (Cyrus, n.d.).
7. This asymmetric relationship between those who collect, store and mine data and their targets is sometimes referred to as the ‘big data divide’ (Andrejevic, 2014).
8. Given the inability, for the individual user, to exploit the economies of scale (‘super-additive insights’) of the data economy, described towards the end of section 3. The data from a single individual is just a drop in the ocean and that converts into a relatively weak economic position.
9. This is the principle requiring that individuals with similar talents and abilities should have the same chances of obtaining positions of authority and responsibility in society (Rawls, 1999, p. 72).
10. This is the principle that inequality should be justified insofar as it is necessary to improve the conditions of the least advantaged citizens. If a more equal distribution is possible that does not worsen conditions for the least advantaged citizens, the existing level of inequality is unjust (Rawls, 1999, p. 72).
11. See also *Analyse and Kritik*, 35(1), special issue on property-owning democracy, with essays from Samuel Freeman, Albert Waele, John E. Roemer, Martin Beckstein, Gavin Kerr, Ivo Walliman-Heimer, Tilo Wesche, Jan Narveson, Jahel Queralt, Michael G. Festl, Andrew Watson, Carina Fourie, Michael Schefczyk, Fabian Schuppert, Emilio Marti, Thad Williamson, Francis Cheneval, edited by Christoph Laszlo and Francis Cheneval (Cheneval & Laszlo, 2013).
12. The gap in question here is the gap between the children of the wealthy, socially networked, and cultured, and the children of the economically, socially, and culturally disadvantaged.
13. In the Rawlsian tradition, health has been identified as such a basic good (Daniels, 1981).
14. The relevance of opportunities provided by control over data assets in the context of political competition relates to the first principle of justice: the ‘equal liberty’ principle. Rawls requires the basic structure of society to satisfy the fair value of political liberties, which is formally analogous to fair equality of opportunity, requiring fair opportunities to influence the political process (Rawls, 1999, p. 197).
15. A rival good is a type of good that may only be possessed or consumed by a single user.
16. However, Art. 20 fails to ensure user-centricity, as it is not clear that the user will be able to decide, for example, to transfer only a specified portion of his or her personal data.
17. An economic regime in which PDPCs play a significant role thus may arise by virtue of combining two social processes. On the one hand, data subjects should exercise their rights to obtain a copy of their data. On the other hand, individuals with a socially-oriented entrepreneurial mindset should be willing to

initiate PDPCs as founders and initial managers, building the infrastructure that enables the collection of personal data by large groups of citizens. University-funded scholars could also play a role by lending their skills to the execution of such projects. We believe that incentives for highly skilled individuals to take this role should be not only economic but also moral. Economic incentives are not excluded *a priori*, since, just like a company, a PDPC may assign to its managers a wage commensurate to their skills once it becomes profitable. However, thanks to the moral and idealistic appeal of the project, one may hope that the economic benefit necessary to motivate highly skilled individuals would be less than what large internet corporations need to spend to attract comparable talent.

18. Notice that control by democratic general assembly is not meant to suggest that members should make difficult decisions about technology and ethics without external aid. A cooperative may, for example, hire ethical and technology experts and take their suggestions into account before voting on specific proposals. Full-time executives are needed to both implement assembly's decisions in practice and submit concrete, feasible proposals to the general assembly's vote. However, the final authority for morally important and self-defining choices rests in the hands of the general assembly, where every member has one vote.
19. Still, unequal chances between persons with different levels of personal interest in the data economy are unavoidable. However, this irrelevant from the standpoint of FEO, since inequalities between persons with different motivations do not count as FEO violations.
20. This claim appeals to Rawls's FEO principle for the economic aspect and to the principle of the fair value of the political liberties (see footnote 14) for the political aspect.

Acknowledgments

The authors wish to thank John Tasioulas and Francis Cheneval for extended discussions of the idea of Property-Owning Democracy and Ian Grant for language improvements and further comments on the paper

Disclosure statement

No potential conflict of interest was reported by Michele Loi.

Ernst Hafen is President of the MIDATA Cooperative. Paul-Olivier Dehaye is director of Geneva-based NGO PersonalData.IO. Both the MIDATA Cooperative and PersonalData.IO are not for profit entities with a stake in the data economy.

Notes on contributors

Michele Loi is a political philosopher by training, Senior Researcher at the Digital Society Initiative of the University of Zurich. His research interest is the ethics of big data, in particular medical data, and algorithms. Michele Loi is the author of 'Giustizia e Genetica' and has published widely, mostly in applied ethics journals (e.g., Journal Of Applied Philosophy, Bioethics, Ethics And Information Technology, Philosophy And Technology). He wrote the first draft of the paper and all revisions.

Paul-Olivier Dehaye is a mathematician, director of Geneva-based NGO PersonalData.IO, focused on personal data rights. He provided feedback and helped revising the paper. @podehaye

Ernst Hafen is Deputy head of Institute for Molecular Systems Biology and President of the MIDATA Cooperative. Ernst Hafen is the author of more than two hundreds contributions in the field of genetics and systems biology. He is among the pioneers of the data cooperative movement and contributed ideas and text to different sections of this paper. @ehafef

References

- Andrejevic, M. (2014). The big data divide. *International Journal of Communication*, 8, 1673–1689.
- Blasimme, A., Vayena, E., & Hafen, E. (2018). Democratizing health research through data cooperatives. *Philosophy & Technology*, 31(3), 473–479. <https://doi.org/10.1007/s13347-018-0320-8>
- Bowles, S. (2005). *Unequal chances: Family background and economic success*. Princeton University Press.
- Brighouse, H., & Swift, A. (2006). Parents' rights and the value of the family. *Ethics*, 117(1), 80–108. <https://doi.org/10.1086/508034>
- Budin-Ljøsne, I., Teare, H. J. A., Kaye, J., Beck, S., Bentzen, H. B., Caenazzo, L., Collett, C., D'Abramo, F., Felzmann, H., Finlay, T., Javaid, M. K., Jones, E., Katić, V., Simpson, A., & Mascalonzi, D. (2017). Dynamic Consent: A potential solution to some of the challenges of modern biomedical research. *BMC Medical Ethics*, 18(1), 4. <https://doi.org/10.1186/s12910-016-0162-9>
- Carballa Smichowski, B. (2016). *Data as a common in the sharing economy: A general policy proposal* [Manuscript submitted for publication]. <https://hal.archives-ouvertes.fr/hal-01386644>
- Cate, F. H. (2006). The failure of fair information practice principles. In J. K. Winn (Ed.), *Consumer protection in the age of the information economy* (pp. 343–379). Ashgate. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1156972
- Cheneval, F. (2018). Property rights of personal data and the financing of pensions. *Critical Review of International Social and Political Philosophy*, 1–23. <https://doi.org/10.1080/13698230.2018.1528521>
- Cheneval, F., & Laszlo, C. (2013). Editorial. *Analyse & Kritik*, 35(1), 1–8. <https://doi.org/10.1515/auk-2013-0119>
- Custers, B., Calders, T., Schermer, B., & Zarsky, T. (Eds.). (2012). *Discrimination and privacy in the information society: Data mining and profiling in large databases (2013 edition)*. Springer.
- Cyran, B. R. (n.d.). *Microsoft ought to kick off search for Bing buyer*. Reuters Blogs. Retrieved April 4, 2016, from <http://blogs.reuters.com/breakingviews/2011/07/22/microsoft-ought-to-kick-off-search-for-bing-buyer/>
- Daniels, N. (1981). Health-care needs and distributive justice. *Philosophy & Public Affairs*, 10(2), 146–179. <https://doi.org/10.2307/2264976>
- Ferretti, T. (2020). A liberal egalitarian perspective on the platform economy: mitigating its distributive effects or changing the organizations running it? *Journal of Social Philosophy*, 51(1), 54–79. <https://doi.org/10.1111/josp.12330>

- Grassegger, V. H., & Krogerus, M. (2016, December 3). Ich habe nur gezeigt, dass es die Bombe gibt. *Das Magazin*. <https://www.dasmagazin.ch/2016/12/03/ich-habe-nur-gezeigt-dass-es-die-bombe-gibt/>
- Hacker, J. S. (2015). *Policy Network—The institutional foundations of middle-class democracy*. https://www.jacobhacker.com/assets/hacker_pn.pdf
- Hartzog, W. (2018). *Privacy's blueprint: The battle to control the design of new technologies*. Harvard University Press.
- Haynes, P., & Carolyn Nguyen, M.-H. (2014). Rebalancing socioeconomic asymmetry in a data-driven economy. In B. Bilbao-Osorio, S. Dutta, & B. Larvin (Eds.), *The global technology report 2014: Rewards and risks of big data* (pp. 67–72). World Economic Forum. http://www3.weforum.org/docs/WEF_GlobalInformationTechnology_Report_2014.pdf
- Hiltzik, M. (2019, April 17). Facebook shareholders are getting fed up with Zuckerberg but can't do anything about him. *Los Angeles Times*. <https://phys.org/news/2019-04-facebook-shareholders-fed-zuckerberg.html>
- Kerr, G. (2015). 'Predistribution', property-owning democracy and land value taxation. *Politics, Philosophy & Economics*, 1470594X15573458. <https://doi.org/10.1177/1470594X15573458>
- Mirsch, T., Lehrer, C., & Jung, R. (2017). Digital nudging: Altering user behavior in digital environments. In *Proceedings Der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017)* (pp. 634–648). <https://wi2017.ch/de/proceedings>
- O'Neil, C. (2017). *Weapons of math destruction how big data increases inequality and threatens democracy*. B/D/W/Y Broadway Books.
- O'Neill, M. (2009). Liberty, equality and property-owning democracy. *Journal of Social Philosophy*, 40(3), 379–396. <https://doi.org/10.1111/j.1467-9833.2009.01458.x>
- OECD. (2013). *Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value* (No. 220; OECD Digital Economy Papers). <https://www.oecd.org/sti/inno/data-driven-innovation-interim-synthesis.pdf>
- OECD. (2014). *Data-Driven innovation for growth and well-being: Interim synthesis report*. <https://www.oecd.org/sti/inno/data-driven-innovation-interim-synthesis.pdf>
- Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
- Rawls, J. (1999). *A theory of justice* (2nd ed.). Harvard University Press.
- Rawls, J. (2001). *Justice as fairness: A restatement*. Harvard University Press.
- Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation 2016/679), European Parliament and the Council. (2016). http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG&toc=OJ:L:2016:119:TOC
- Right to data portability, 20, European Parliament, EU General Data Protection Regulation (EU-GDPR). (2016). http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf
- Savage, M., Devine, F., Friedman, S., Laurison, D., McKenzie, L., Miles, A., Snee, H., & Wakeling, P. (2015). *Social Class in the 21st Century*. Penguin.
- Solove, D. J. (2012). Introduction: Privacy self-management and the consent dilemma. *Harvard Law Review*, 126, 1880.
- Weinmann, M., Schneider, C., & Vom Brocke, J. (2016). Digital Nudging. *Business & Information Systems Engineering*, 58(6), 433–436. <https://doi.org/10.1007/s12599-016-0453-1>
- World Economic Forum, & Bain & Company, Inc. (2011). *Personal data: The emergence of a new asset class*. World Economic Forum.