The future of data ownership: An uncommon research agenda

Jacqueline Hicks
School of Politics and International Relations, University of Nottingham, UK

Abstract

Responding to the current data monopoly by Big Tech firms, there is increasing interest in the potential for collective ownership of data in a 'data commons'. This article aims to introduce the topic to non-specialists, highlight the broader social significance of data ownership, and reframe the data commons as a problem of social interaction rather than regulatory design. It critically engages with existing work on the data commons to suggest alternative directions for research which draw on humanities and interpretive social science approaches. It recommends the addition of research that (1) includes knowledge from non-Western legal traditions to overcome objections to the possibility of data ownership, (2) studies existing social relations of data governance rather than seeks to design ideal regulations, (3) prioritises a focus on the social and communicative practices involved in a data commons rather than seeing it as a solution to a specific problem, and (4) analyses the meaning-making practices of people as they engage in their own data governance, and the evolving narrative around what constitutes a data commons. The purpose of this alternative research agenda is to provide discursive scaffolding for the ongoing regulatory work on the data commons, while ensuring that its full potential remains imaginatively open.

Introduction

Many internet users are seduced by the digital services of Big Tech companies like Google, Facebook and Amazon, while remaining distrustful of their business practices. The ubiquitous unease with these companies' monopolisation of our personal data is driving two very different types of response.

The first type of response uses or extends *existing remedies*, like anti-monopoly laws or trade regulation orders. Amid continuing speculation about US government action to break up the largest tech companies (Brent & McKinnon, 2019), 2019 saw one of the

Corresponding author:
Jacqueline Hicks, School of Politics and International Relations, University of Nottingham, Nottingham, NG7 2RD, UK.
Email: j.hicks1@ids.ac.uk

largest penalties ever imposed by the US Federal Trade Commission against Facebook for privacy violations (Federal Trade Commission, 2019). In 2018, the EU introduced the General Data Protection Regulation (GDPR) giving European internet users more rights to request their data from online companies, clearer information about how their data are processed, and giant fines for companies that do not comply. The GDPR kicked off a surge in new data protection regulation around the world (Schwartz, 2019).

Other suggestions offer ways for users to gain a slice of the pie for themselves. The Governor of California recently proposed a 'data dividend' for tech companies like Facebook to pay its users a percentage of the revenue derived from their data (Au-Yeung, 2019). Third party platforms, like 'digi.id' and 'datacoup', offer internet users a way to monetise their own data (Elvy, 2017). What all of these efforts have in common is that they accept the basic 'data for access' bargain between users and tech companies, but seek to either dis-incentivise potential harm from data use or raise compensation for users.

The second type of solution to Big Tech data monopolisation is to develop *alternative systems* of value creation from personal data through some form of collective or communal data ownership. They are alternative systems of value creation because they do not necessarily involve monetary gain, although they could. A non-monetary transaction could involve individuals donating their data to a worthy research project, while a monetary transaction could allow people to sell their data. They are collective or communal systems because they imply some form of decentralised management where those who contribute their data also have some input into collective decision-making about how to use the data.

The interest in alternative systems of value creation from data now coincides with a post-Covid appetite for a radical re-think of economic systems. However, scholars have already been highlighting the broader socio-economic significance of data economics for a number of years. In a measure of data's importance to the organisation of society, Karl Polanyi's *The Great Transformation* (1944) has experienced something of a resurgence in the scholarly literature. Particularly for those writing critically about Big Tech data monopolisation, Polanyi's analysis of how industrial capitalism commodified land, labour and money provides a useful template for understanding the implications of today's use of data (Cohen, 2019; Couldry & Mejias, 2019; Zuboff, 2019). While Zuboff (2019) meticulously folds the exploitation of personal data into a production process of 'surveillance capitalism', Couldry and Mejias (2019) elevate its significance into the development of a new 'social order of datafication' which shares the scope and scale of colonialism.

As an antidote to these dystopian visions, the idea of communal data ownership has captured the imagination of a broad range of people – journalists, technologists, academics, third sector researchers, and governments in a truly interdisciplinary space. It is both an exciting and challenging research object because it has so many different elements – economic, legal, technical and sociological – that need to be understood together in a fast-changing environment. It requires a research agenda that not only draws from different fields, but also from different types of knowledge, epistemologies and approaches.

Extant literature has increased our understanding of specific aspects of a data commons, making progress on suggesting ways to design ethical, regulatory and legal regimes. But many of the writings are bound by the conventions of the fields they are

associated with. What research directions on the data commons are available to anthropologists, social theorists, or interpretive social scientists?

This article seeks to further research on communal data ownership by developing some research approaches for those who are more interested in social organisation than in data per se. While it engages with some of the existing debates on communal data ownership, it will foreground some of the assumptions that accompany the legal and technical approaches used in the literature. It then shows some of the additional insights that can be gained on this topic by shifting the research agenda onto a more exploratory experimental terrain. It has the ambition to draw interest among non-specialists by explaining some key novel features of data as a resource, highlighting the broader social significance of data as property, and reframing the data commons as a problem of social interaction rather than regulatory design.

The article begins by laying out some of the basics about communal data ownership. It briefly highlights three unique features of data as a resource, and gives some examples of existing and proposed communal data management. Prefaced by a short consideration of the societal significance of property rights, the article then reviews some of the current academic literature on the possibility of property rights for data, and by extension a data commons. Rather than engaging in these arguments on their own terms, the article suggests expanding the research agenda with four alternative complementary approaches to exploring communal data ownership.

In sum, the four alternative research directions suggested are:

- The inclusion of non-Western legal traditions as sources of knowledge about communal ownership rather than the market-oriented, individualistic approach to property found in Western law.
- The study of existing social relations of data governance rather than suggesting ideal regulatory designs.
- Prioritising the study of communicative processes within data commons research projects rather than their ability to solve a particular problem.
- The analysis of the narrative, rhetoric and meaning-making that exist in writings and practices associated with a data commons.

Three unique features of data as a resource

Compared to other types of 'resource', data have specific properties which make them both especially suitable for communal ownership and difficult to achieve. This section uses metaphor to highlight three material features of data that differentiate them from other types of 'resource'.

Data are ambiguous

I describe data as ambiguous because they defy definitive categorisation in two main ways. First, data are frequently described as either 'personal data', 'open data' or a mixture of the two. A widely used definition of personal data is the EU's 'any information relating to an identified or identifiable natural person' (European Data Protection Supervisor, undated).

Open data refer to data which are open for free access, use and modification to be shared for any purpose, such as data about traffic flows or weather. However, there are also ambiguous borderline cases between personal and open data, for example on energy consumption in people's houses (Smichowski, 2018).

Second, personal data in particular have multiple values. They have economic value, under certain circumstances, as we will see below. However, unlike other types of resource, data also have non-economic value. For example, at societal level data could help control the spread of a disease; for the individual, they have privacy value; for the state, value as material for surveillance and security. Such confluence of values makes the commercial exploitation of personal data ethically more complex compared to other types of economic resource.

Data are magnetic

I use the metaphor of magnetism to describe two features of personal data about how they combine. The first point is that they tend to be about multiple people rather than individuals by the nature of the processes that create data (Metcalf & Crawford, 2016; Taylor, 2006). Most of the data about us are also about other people – our bank accounts reveal details about those we transact with, our emails record other people's replies (Open Data Institute [ODI], 2018). Even if we deny consent, or do not use data-generating devices, an organisation can use data about other people who share similar characteristics, like postcode, gender or age, to make statistical extrapolations that will later affect you in some way (Purtova, 2017; Tisne, 2018). In short, the networked aspect of personal data coupled with the social aspect of our lives means a lot of personal data are communal by nature (Roessler & Mokrosinska, 2013).

The second magnetic feature is that value is created from personal data when they are analysed in combination with other data sets. This can be a combination of multiple data sets about one person to give an ever more complex picture of a person's character for advertising – a '360 profile'. This can then be analysed with similar data from thousands of other people to 'bucketize' you, like 'heavy smoker with a drink habit' or 'healthy runner, always on time' (Tisne, 2018). Different data sets about groups of people can be brought together for all sorts of insights, the results of which then become valuable as input for some other prediction or inference.

Data are abundant

Data can also be described as 'abundant' because they can be copied and used many different times for different purposes compared to other types of economic resources like oil or land which are depleted when used. In legal parlance this is known as 'non-rivalrous', meaning that data's use does not reduce their availability to others. Others have described this abundance as 'multiple' (Prainsack, 2019b).

This feature of data is somewhat mitigated by the way that data gain value, since other inputs have to be combined with data to make them valuable. Data analysts typically spend a lot of time and effort cleaning data, as well as analysing them and matching them to a problem that needs to be solved, like how to connect car owners who want to earn

extra money with people who need a lift. Carballa Smichowski (2018) elegantly describes it thus: 'The value created from data is the combined result of the dataset and the analysis applied to it. The two are inherently inseparable in the value creation process, just like the movement of a car cannot be fully attributed either to the engine or to fuel' (p. 20).

Even though data need some inputs to make them valuable, the fact that they can be copied over and over again draws them into radical reappraisals of economic organisation. If data are not wholly subject to the same natural limitations as other types of resource like land or oil, then they have the potential to contribute to what some call an 'economy of abundance' (P2PFoundation, 2019). This concept is used in contradistinction to the 'economics of scarcity' that characterise the current economic system where value is created by erecting legal barriers like copyright to produce false scarcity. Needless to say, this is a contested idea (see Peters, 2020).

Imagined and existing forms of data commons

Currently there is a lot of heat in the idea of a data commons, with a large and fast-growing body of academic, popular media and grey literature all highlighting its potential. They include suggestions of 'data trusts' (e.g. Blankertz, 2020; Delacroix & Lawrence, 2019), 'data cooperatives' (Ada Lovelace Institute [ADI] & AI Council, 2021), 'data coalitions' (Prewitt, 2021; The Data Freedom Act, 2020), 'data unions' (Arrieta-Ibarra et al., 2018), as well as 'data commons' (Prainsack, 2019a; Yakowitz, 2011).

There are no agreed definitions of any of these terms beyond the general idea of collective data management by non-state, non-market actors, but efforts are underway to clarify ways to differentiate between them (Micheli et al., 2020). Broadly, they are imagined on a continuum between a more centralised internal authority which make decisions about data use on behalf of their members, and a more decentralised participatory decision-making structure. They draw vaguely on their wider associations, so that coalitions, cooperatives and commons imply more decentralised decision-making, while trusts imply the appointment of a trustee. Blankertz (2020) envisages data trusts increasing the bargaining power of individuals by negotiating with businesses about the conditions of data use. In more general terms, common property implies that the resource in question is available for use by its members, while collectives refer only to the process of decision-making (Waldron, 1988 in Hummel et al., 2020).

While much of the writing on communal data ownership remains speculative, there are already examples of its practice in specific domains. The particular benefits of pooling health-related data have driven data commons platforms like CommonHealth where individuals deposit their personal health data and decide how to share them, or the BioBank, which regularly collects the health data of 500,000 participants to share with researchers. More controversially, Google's parent company Alphabet got three years into actively planning a smart city section of Toronto fitted out with sensors for traffic, building and energy usage. A 'civic data trust' was planned and public consultations held to gauge responses to data governance proposals including the collection of health, personal mobile and consumer data. In the face of opposition, the whole plan was abandoned in 2020 (Koetsier, 2020). A more grassroots-led data commons vision was initially tried in Barcelona until 2019. In practice, it was less encompassing and so less ambitious

than Google's abandoned experiment in Toronto, but it served the crucial function of 'opening up a new policy-data interaction' for its citizens (Calzada & Almirall, 2020).

The significance of property rights

The communal ownership of data, just like the communal ownership of land or any other resource, is a type of property rights regime. Property rights are highly significant as one of the connective elements that bridge forms of economic, social and political organisation. In the broadest sense, property rights are cornerstones of society, structuring the different ways that wealth can be acquired, used and transferred as societal values change and develop (Benda-Beckmann et al., 2011). The property regime we take for granted today that underpins industrial capitalism took centuries to develop, and this process is only just beginning for data capitalism.

At present, the data economy rests on shifting legal sands as governments and wider societies debate the types of rights, controls and responsibilities to surround the collection and use of personal data. In place of a settled agreement, Big Tech firms currently use consent and permissions procedures to appropriate the data they collect in exchange for access to their services. These procedures are seen by some as largely meaningless as their length and complexity preclude any real conscious rational and autonomous choice about how our personal data are processed (Schermer et al., 2014, p. 171).

In effect, what we are experiencing now is the real-time development of property rights over personal data. It is the beginning of what will be a long regulatory journey, and the decisions being made now are setting the pathways for future economic, social and political configurations.

Existing literature on property rights for data

In the academic literature, scholars debate whether data *should* be propertised (given the value to society and individual privacy), and whether data *can* be propertised (given data's 'magnetic' and 'abundant' features).

Legal scholars have been rehearsing arguments about the legal logic and ethics of applying property rights to data for decades already, comprehensively summarised in Purtova (2011). At the heart of objections to the idea is a concern that propertisation is synonymous with commodification, further enabling data markets. This is considered undesirable given the intensely personal nature of data about us, and the power asymmetries that already exist between data subjects (us) and data collectors (Big Tech).

Hummel et al. (2020) summarise some scholarly concerns against applying property rights to data, including the loss of privacy that would result after a person's data have been commercialised (Montgomery, 2017), unrealistic expectations about how much a person's data could be worth if sold Evans, 2011), and that selling your own data could undermine an individual's 'personhood' because people 'do not just own information; they are constituted by it' (Floridi, 2014 in Hummel et al., 2020, p. 558). While Lessig (2002) notes that property rights over data are often resisted because they are thought to isolate individuals.

The second point about whether data *can* be propertised has spawned an ongoing debate related to the unique features of data. As we have seen, unlike other resources, the possession of data does not imply that one is the sole possessor and exclusive user. In other words, several people can use data at one and the same time, and it is almost impossible to exclude third parties (Hummel et al., 2020). Practical problems of ownership also arise from the difficulty of distinguishing one person's data from another, and separating personally identifiable data from other types of data.

Purtova (2017, p. 77) summarises these difficulties: 'the same piece of data may behave as personal and non-personal under different circumstances: it may be more or less identifiable, and have a stronger or weaker link to a person, and the moment of transition from one state to another may pass unnoticed by the data holder and the affected individuals and groups'. Similarly, Taylor et al. (2017) point out that the boundaries for group ownership of personal data may not be any easier to define than for individuals.

There is also disagreement about how compatible current legal frameworks are with the idea of data ownership. Prainsack (2019a) detects a general schism between US law, which allows for some of the rights associated with ownership for data, while European legal traditions treat personal data as a human right rather than potential property. Hummel et al. (2020) summarise the different opinions as those who see data ownership as incompatible with the law, those who say the law is unclear, and those who say that some elements of data ownership already exist in law.

In view of these ethical, technical and legal issues, some have argued to drop the idea of data ownership in favour of a 'bill of data rights'. Tisne (2018) makes the case persuasively, and the idea has filtered through to Members of the British Parliament (Byrne, 2018), various international think tanks and policy organisations (ODI, 2019; Ubaldi et al., 2019; Urban, 2019), and the academic literature (Berman & Hirschman, 2018). A bill of data rights would give protection about what sort of data can be collected, and how they are used or shared after collection, and would be easier to establish than any attempts to create property rights. Others suggest hybrid regimes, like Smichowski's (2016) idea of combining 'reciprocity licenses' with a data commons, or Prainsack's (2019a) support for a form of data commons that pays particular attention to practices of inclusion and exclusion. McMahon et al. (2020) propose institutions called 'Harm Mitigation Bodies' to operate while other legal remedies are being designed, and Hummel et al. (2020) conclude that a form of 'quasi-ownership' lying somewhere between individual and collective ownership could be the answer depending on a societal debate about the relative balance between the two poles.

The array of different regulatory propositions is testament to both the importance of the topic and the dense, detailed thinking that is required to suggest new forms of more distributed data governance and ownership. The strong disagreements throughout this literature about how to proceed may seem to have little in common with each other, but from another perspective they share the same overall approach of suggesting ways to design ethical, regulatory and legal regimes. This is not the only way to do social science.

Expanding the research agenda

The very real difficulties of ascribing some form of ownership rights over personal data cannot be wished away, but nor can the very deep significance of ownership if people are

to claim their own stake in the data economy. Rather than trying to shoehorn data ownership into existing legal frameworks, some scholars suggest now might be the right time for a more fundamental re-think. Legal scholar Frank Pasquale argues that the relative plasticity of the data economy compared to industrial capitalism calls for deeper discussions:

> As more economic value is located in software systems, 'big data', pattern recognition, and the 'lords of the cloud' with privileged access to all these processes, we ought to feel more free to reimagine the terms of social cooperation – not less. (Pasquale, 2014, p. 36)

Anthropological approaches are well-suited to this task with their emphasis on rolling back the top layer of regulatory practices to reveal their underlying social functions. Interpretive forms of sociology that focus on the meaning people give to their own actions also have an important role to play in supporting these 'deeper discussions'.

The next sections draw on scholarship about communal ownership from these traditions to offer additional possibilities for research on the data commons. The four directions suggested below are meant as complementary approaches to the extant literature as described above in a shared effort to help 'move the ball down the field' in ongoing work on the communal ownership of data.

Learn from non-Western legal traditions

One issue with many contemporary discussions of data ownership is their narrow definitions of property rights. If the concept of property rights is opened up to reveal the 'heavy freight of political and ideological baggage' it is loaded down with (Benda-Beckmann et al., 2011, p. 2), it becomes apparent that many of the objections to data ownership are actually objections to property rights *as they are currently conceived in the West.*

> Thinking about law without thinking about where it comes from and about all of the things that make it important and interesting is dreadfully dull . . . [It is like] thinking about something that only has meaning in relation to something else, without thinking about the something else. (Griffiths, 2006, pp. 67–68)

Taking a wider view of property, both historical and international comparisons show the many different types of property regimes which are not centred on the individual. This then undermines the assumption that property rights are not appropriate for personal data because data always concern multiple people rather than individuals. While legal entitlements to resource production and consumption are commonly ascribed to *individuals* in Western legal systems (Madison et al., 2018, p. 664), proprietary rights within customary law in Africa and Asia more commonly allow a number of specific interests to be vested in *different holders* (Bennett, 2006, p. 654). As Pierson (2013, p. 96) points out, the common assumption that property is owned by individuals is the politically constructed result of the development of industrial capitalism.

This wider view of property rights also undermines the ethical assumption that the granting of property rights necessarily facilitates market exchange. In other legal systems around the world, some of the main functions of property rights are to protect rights of

use, or to support particular social obligations (Bennett, 2006, p. 654). Indeed, even the notion of property rights is sometimes said to convey a selective and specifically Western legal character to property relationships. Property in the most general sense concerns the ways in which the relations between society's members with respect to valuables are given form and significance (Benda-Beckmann et al., 2011, p. 9).

Other non-Western legal practices also have particular relevance for the development of a data commons. It is no coincidence that most of the research on actually functioning collective property rights regimes for natural resources has taken place in the global South where they tend to occupy larger spaces than in the West. This was certainly the case for the research associated with Eleanor Ostrom, the Nobel Laureate whose work challenged the fallacy of the 'tragedy of the commons' by seeking to understand how communities handle collective problems of natural resource use, like woodlands or fisheries.

Ostrom's research found that an overriding feature of existing communal property arrangements for natural resources is the existence of 'polycentric systems of governance' characterised by multiple governing authorities at different levels. While formally independent of each other, they nevertheless enter into contractual and cooperative undertakings, with recourse to central mechanisms to resolve conflicts (Ostrom et al., 1961, p. 831).

These local units of polycentric governance in communal property arrangements often sit within national systems of 'legal pluralisms' frequently found in developing countries. This is where multiple legal systems operate within one country, often a result of the imposition of colonial law in a territory that had, and continues to have, its own legal system.

The concept and practice of polycentric governance and perhaps even legal pluralisms has relevance for a data commons because there will be no single perfect institutional configuration ready to be deployed, no matter how much thought is put into it. In a clear analysis of some of the different alternative data governance arrangements on offer, Smichowski (2019) concludes that there is no 'one-size-fits all' model, and that a workable alternative data ecosystem can only be built on a variety of data governance models. The current dominance of a handful of Big Tech firms required massive injections of venture capital during the initial years of financial loss before user numbers increased enough to allow the companies to turn a profit. Any form of data commons will not be afforded that same luxury, and so the field will be much more fragmented.

Polycentric governance is therefore likely to proliferate alongside the growth of alternative forms of data ownership. Given the transnational nature of current personal data protections, which apply to a person's data in whichever country the data are stored or used, data governance is also subject to a plurality of legal regimes. In the socio-legal perspective, such pluralism is seen less as an obstacle to the creation of a sound legal order, and more as an inherent trait of an evolving legal order (Zumbansen, 2010), with different systems operating together in what one eminent anthropologist described as a 'working misunderstanding' (Geertz, 1983, p. 231).

As research on data governance becomes more internationalised (Milan & Trere, 2019), locating experimental research in data commons to developing countries where there are existing social practices around the communal governance of natural resources could then be one interesting avenue for research. By following through on their framing of datafication as a digital colonialism, Couldry and Meijas (2019) call for a 'decolonial

approach to data' that reclaims elements erased by Eurocentrism to frame possible 'counterpresents' (p. 111). Exploring approaches to managing data commons from the global South could be one way to do that, not as additional inclusion, but as a lead.

Shift the analytical gaze from desired to existing practices

The bulk of the existing work on data ownership comes from legal scholars and technologists who, by the nature of their work, prefer formal regularity and consistency to any kind of contingency, shapelessness or obfuscation. Where social scientists contribute, they tend to take an economistic approach, understanding the significance of property to be in how it manages the social and economic effects of scarcity and efficiently satisfies human needs (Benda-Beckmann et al., 2011). However, scholarship on property does not have to be so instrumental and normative, it is also possible to approach it as an empirical, descriptive and analytical task (Benda-Beckmann et al., 2011). It is this perspective that is currently missing from conversations about property rights in personal data.

Benda-Beckmann et al. (2011) usefully provide an analytical framework that incorporates the several distinct analytical layers at which property manifests: not just in legal systems, but also in ideologies, actual social relationships and social practices. Exhorting us to study the interrelations between these phenomena, the authors recommend moving the analytical gaze away from desired (just, efficient) states of ideal property relationships to accurate descriptions and explanations of existing property regimes. In practice, this means having the freedom to conduct empirical research that does not necessarily translate directly into regulatory design.

Again, Ostrom's prodigious body of work on the communal ownership of natural resources resonates with this approach. Many other scholars have already recognised the relevance of Ostrom's work for a data commons. Her research group developed a clear framework to explore the formal and informal rules around communal governance that enabled communities to avoid overuse. The framework, resulting in eight design principles, is both detailed and flexible enough to provide a useful way to make sense of the very wide range of communal arrangements her research encountered. It also makes it particularly relevant for thinking about data as a resource with many different forms and uses. Some have usefully employed Ostrom's design principles to suggest various designs for a data commons and/or clarify the differences between commons arrangements for material and non-material resources (Fisher & Fortmann, 2010; Fuster et al., 2017; Mills, 2019; Šestáková & Plichtová, 2019). Indeed Ostrom herself turned to the issue shortly before her death (Hess & Ostrom, 2007).

However, a principle of Ostrom's work which is not so frequently cited in the literature on the data commons is her scepticism of top-down designs and solutions.

> Bureaucrats sometimes do not have the correct information, while citizens and users of resources do. (Ostrom in Ringstrom & Vinocur, 2009)

Ostrom railed against regulatory panaceas based on universal models (Ostrom, 2007), whether they are based on private ownership, state action or indeed communal ownership regimes. Her research pointed to the wisdom of resource users themselves, saying

they often know best how to organise and govern resources. Although they sometimes fail, such examples were as important to study as the ones which succeeded. Focusing less on suggesting regulatory solutions, and more on understanding what works and what does not in real-world situations meant systematic observation of communities in multiple case studies from around the world.

Fieldwork that focuses on communicative processes not solutions

How does systematic observation of communities in a data commons work in practice?

There are already data commons projects, most of which are tightly focused on solving a specific problem. For example, a sleep data commons, where data from different researchers working on sleep disorders are gathered and made accessible to other researchers (Zhang et al., 2018). Agricultural data are another area for building a data commons, with the aim of improving food security (Baarbb & de Beer, 2017). These are wonderfully useful initiatives with the potential to demonstrate real impact.

This article suggests a different focus of research – one that deals with the types of personal data that are generated from day to day online activities and cause general anxiety about privacy intrusions by Big Tech companies. It also suggests a focus on the communicative processes involved in communal data management rather than the degree to which a data commons can help solve a particular problem.

One way to undertake such empirical fieldwork is by deploying relatively inexpensive technologies which store copies of all the data we generate onto small personal servers which are located in the home. One academic project which developed such a system is Databox (Haddadi et al., 2015). Everyone who has a 'databox' can store their own personal data generated from activities such as online search, social media use, personal devices like fitbit or home smart devices like fridges. Companies or organisations that want to use these personal data have to request access.

Databox is part of a wider trend to 're-decentralise' the internet. This is the effort to offer alternatives to the current technical structure of the internet, which relies on a relatively small number of servers that store users' personal data as they engage with Big Tech services. A decentralised internet structure would rely instead on networks of personal servers from a large community of users. The original creator of the World Wide Web, Sir Tim Berners-Lee, has a project called Solid which is building the infrastructure for individuals to have their own personal online data stores – 'Pods' in their homes. Many other projects, groups and organisations are involved in similar enterprises, such as Blockstack, Sandstorm and Dweb, to name just a few. One, called Urbit, has a governance structure drawn from alt-right philosophy (Smith & Burrows, 2021).

Within the context of this article, the point is that technologies such as Databox enable the observation of user communities in a data commons containing much broader types of personal data than usual. Already existing platforms, such as those mentioned above, are other possible sites for fieldwork.

The other key shift suggested in this article is a focus on the communicative processes involved in communal data management rather than how far it can help solve a particular problem. Drawing again from Ostrom's work on the communal management of natural resources, she found that understanding the communicative processes of 'cooperation

without external enforcement' was key to the success of commons arrangements. Those that failed contained groups of people who had 'no capacity to communicate with one another, no way to develop trust, and no sense that they must share a common future' (Ostrom, 1990, p. 21).

It is not just a shift in focus to communication that is important here, but also to process. This is what is captured in Linebaugh's (2009) concept of 'commoning' – an actionable verb that moves beyond analysis of 'a commons' to the dynamic practices used by communities to co-manage resources (Stavrides & De Angelis, 2010, in Katrini, 2018).

Madison et al. (2018) helpfully provide some examples of relevant research questions: What sort of decisions would people make about their personal data if given the chance? Who is a member, and who decides who may be a member? How are data contribution and extraction monitored and, if necessary, limited? What sanctions and dispute resolution mechanisms are provided for misconduct? To what extent do these self-governance mechanisms rely on or incorporate formal legal mechanisms, and to what extent do they rely on or incorporate other, non-legal institutions or social structures?

Examine the narrative and rhetoric associated with data commons

The recent increase in writings about the data commons, its indeterminacy and speculative, future-facing character make it a good candidate for examining the narratives being built around it. This can include straightforward identification of different types of data commons, comparing for example the European Commission's idea of a data commons for industrial data with more citizen-centric notions associated with open data in cities (Calzada & Almirall, 2020). What is missing from the literature is the more reflective type of analysis that is associated with the interpretive social sciences.

Blue (2016, p. 68) summarises the interpretive approach as recognising 'that humans are interpretive creatures who often disagree about meanings, values, means and ends; that knowledge and meaning-making practices inform and are informed by the social contexts in which they are located; and that social relations are constituted in important ways by arrangements of identity, knowledge and power'. Examining the way that knowledge about the communal ownership of data is produced and elaborated could be an interesting addition to scholarship on the topic.

A rich seam of work already exists on the narratives around an 'information commons', which seeks to rebalance the entitlements of users and producers of creative works through the copyright regime. For example, Vaidhyanathan (2003) notes the rhetorical value of activists using 'property talk' and 'commons talk' in relation to intellectual property. Similarly, Lessig (2002) notes that data do not receive the same amount of protection as copyright affords to creative works because the 'social meaning' of privacy is not constructed as property in the same way as books or music. Madison et al. suggest the utility of examining the 'cultural commons' in 'expressive terms rather than purely functional terms, looking to the construction and evolution of meaning in the system, as reflected in symbol and narrative' (2018, p. 673).

In the context of an empirical research project to experiment with setting up a commons of users' personal data, detailing how people understand their own actions as they participate in efforts at governance would be an important insight to the topic as it evolves.

Conclusion

This article has argued for research on the data commons to include sociological experiments where people collectively manage their own personal data from their online trails, social media and smart devices. It suggests research that focuses on people's actual communicative processes and reflexive meaning-making practices, rather than seeking optimal legal or ethical design to solve specific problems with specific types of data. The aim is not to dismiss the crucial work of legal scholars, technologists and social scientists designing regulatory solutions for data commons, but to support them with some conceptual and discursive scaffolding.

Placing the social relationships that people create or maintain at the heart of research on data commons could also help progress the debate on property rights over data. Legal scholars note that over the long term, law can develop from social practices ('private ordering') through to a more accepted normative framework ('customary law'), to be either taken up or ignored by the state ('public ordering') (Bennett, 2006, p. 642; Deakin, 2015). It may not be possible to neatly assign property rights to data as we currently understand them, but change in the meanings of property and ownership may already be under way. Personal data as a new object of value may stretch the bounds of property definitions and categories, so the idea is to study the concept of property as it transforms, rather than abandoning it. Learning from the way polycentric governance and legal multiplicities operate in the non-Western world may help to navigate these changes.

As personal data become implicated in all sorts of economic and social injustices, there is a quite natural sense of urgency to pin down their regulation in mitigation. However, the current and near future reality is colliding multiple regulatory systems at local, national and international levels. Under these circumstances, humanities methods are well-suited to the issue in their non-teleological embrace of ambiguity and uncertainty. The question which guides research on data commons is a normative one: Who should have the right to use personal data, and how should the value derived from the data be shared? In the humanities tradition, part of the answer lies in provoking new conversations and interpreting existing ones. It may also serve to inspire public engagement in a way that questions of individual privacy have so far failed to do.

References

Ada Lovelace Institute & AI Council. (2021). *Exploring legal mechanisms for data stewardship*. www.adalovelaceinstitute.org/feature/data-cooperatives/

Arrieta-Ibarra, I., Goff, L., Jiménez-Hernández, D., Lanier, J., & Weyl, E. G. (2018). Should we treat data as labor? Moving beyond 'free'. *AEA Papers and Proceedings*, *108*, 38–42. https://doi.org/10.1257/pandp.20181003

Au-Yeung, A. (2019, February 14). California wants to copy Alaska and pay people a 'data dividend.' Is it realistic? *Forbes*. www.forbes.com/sites/angelauyeung/2019/02/14/california-wants-to-copy-alaska-and-pay-people-a-data-dividend–is-it-realistic/#4af53eda222c

Baarbb, J., & de Beer, J. (2017). A data commons for food security. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3008736

Benda-Beckmann, F., Benda-Beckmann, K., & Wiber, M. (2011). The properties of property. In K. Benda-Beckmann, F. Benda-Beckmann, & M. Wiber (Eds.), *Changing properties of property* (pp. 1–40). Berghahn Books.

Bennett, T. (2006). Comparative law and African customary law. In M. Reimann & R. Zimmermann (Eds.), *The Oxford handbook of comparative law* (pp. 641–673). Oxford University Press.

Berman, E. P., & Hirschman, D. (2018). The sociology of quantification: Where are we now? *Contemporary Sociology: A Journal of Reviews*, *47*(3), 257–266. https://doi.org/10.31235/osf.io/4sgc2

Blankertz, A. (2020). *Designing data trusts: Why we need to test consumer data trusts now*. Stiftung Neue Verantwortung. www.stiftung-nv.de/en/publication/designing-data-trusts-why-we-need-test-consumer-data-trusts-now

Blue, G. (2016). Framing climate change for public deliberation: What role for interpretive social sciences and humanities?, *Journal of Environmental Policy & Planning*, *18*(1), 67–84. https://doi.org/10.1080/1523908X.2015.1053107

Brent, K., & McKinnon, J. (2019, June 3). Congress, enforcement agencies target tech. *Wall Street Journal*. www.wsj.com/articles/ftc-to-examine-how-facebook-s-practices-affect-digital-competition-11559576731?mod=article_inline

Byrne, L. (2018, August 30). *Bill of data rights – speech in committee stage of the Data Protection Bill*. https://liambyrne.co.uk/bill-of-data-rights-speech-in-committee-stage-of-the-data-protection-bill/

Calzada, I., & Almirall, E. (2020). Data ecosystems for protecting European citizens' digital rights. *Transforming Government: People, Process and Policy*, *14*(2), 133–147. https://doi.org/10.1108/tg-03-2020-0047

Cohen, J. (2019). *Between truth and power: The legal constructions of informational capitalism*. Oxford University Press.

Couldry, N., & Mejias, U. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.

Deakin, S. (2015). *Law as evolution, evolution as social order: Common law method reconsidered*. Working Papers wp470, Centre for Business Research, University of Cambridge. https://doi.org/10.2139/ssrn.2640917

Delacroix, S., & Lawrence, N. (2019). Bottom-up data trusts: Disturbing the 'one size fits all' approach to data governance. *International Data Privacy Law*, *9*(4), 236–252. https://doi.org/10.1093/idpl/ipz014

Elvy, S. (2017). Paying for privacy and the personal data economy. *Columbia Law Review*, *117*(6), 1369–1459.

European Data Protection Supervisor. (undated). *Personal data*. https://edps.europa.eu/data-protection/data-protection/glossary/p_en

Evans, B. (2011). Much ado about data ownership. *Harvard Journal of Law & Technology*, *25*(1), 69–130.

Federal Trade Commission. (2019, July 24). *FTC imposes $5 billion penalty and sweeping new privacy restrictions on Facebook*. www.ftc.gov/news-events/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions.

Fisher, J., & Fortmann, L. (2010). Governing the data commons: Policy, practice, and the advancement of science. *Information & Management*, *47*(4), 237–245. https://doi.org/10.1016/j.im.2010.04.001

Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.

Fuster, M., Smichowski, B., Smorto, G., Espelt Rodrigo, R., Imperatore, P., Rebordosa, M., Rocas, M., Rodríguez, N., Senabre, E., & Ciurcina, M. (2017). *Multidisciplinary framework on commons collaborative economy*. DECODE, UOC. www.decodeproject.eu/publications/multidisciplinary-framework-commons-collaborative-economy

Geertz, C. (1983). Local knowledge: Fact and law in comparative perspective. In Geertz, C. *Local knowledge: Further essays in interpretive anthropology* (pp. 167–234). Basic Books.

Griffiths, J. (2006). The idea of sociology of law and its relation to law and to sociology. In M. Freeman (Ed.), *Law and sociology: Current legal issues* (pp. 49–68). Oxford University Press.

Haddadi, H., Howard, H., Chaudhry, A., Crowcroft, J., Madhavapeddy, A., McAuley, D., & Mortier, R. (2015). Personal data: Thinking inside the box. *Aarhus Series on Human Centered Computing*, *1*(1). https://doi.org/10.7146/aahcc.v1i1.21312

Hess, C., & Ostrom, E. (2007). Introduction: An overview of the knowledge commons. In C. Hess, & E. Ostrom (Eds.), *Understanding knowledge as a commons: From theory to practice* (pp. 3–26). MIT Press.

Hummel, P., Braun, M., & Dabrock, P. (2020). Own data? Ethical reflections on data ownership. *Philosophy and Technology*, *34*, 545–572. https://doi.org/10.1007/s13347-020-00404-9

Katrini, E. (2018). Sharing culture: On definitions, values, and emergence. *The Sociological Review*, *66*(2), 425–446. https://doi.org/10.1177/0038026118758550

Koetsier, J. (2020, May 13). 9 things we lost when Google cancelled its smart cities project in Toronto. *Forbes*. www.forbes.com/sites/johnkoetsier/2020/05/13/9-things-we-lost-when-google-canceled-its-smart-cities-project-in-toronto/

Lessig, L. (2002). Privacy as property. *Social Research: An International Quarterly*, *69*(1), 247–269.

Linebaugh, P. (2009). *The Magna Carta manifesto: Liberties and commons for all*. University of California Press.

Madison, M. J., Frischmann, B. M., & Strandburg, K. J. (2018). Constructing commons in the cultural environment. *Cornell Law Review*, *95*(4). https://doi.org/10.31228/osf.io/76g93

McMahon, A., Buyx, A., & Prainsack, B. (2020). Big Data governance needs more collective responsibility: The role of harm mitigation in the governance of data use in medicine and beyond. *Medical Law Review*, *28*(1), 155–182. https://doi.org/10.1093/medlaw/fwz016

Metcalf, J., & Crawford, K. (2016). Where are human subjects in Big Data research? The emerging ethics divide. *Big Data & Society*, *3*(1). https://doi.org/10.1177/2053951716650211

Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. *Big Data & Society*, *7*(2). https://doi.org/10.1177/2053951720948087

Milan, S., & Trere, E. (2019). Big data from the South(s): Beyond data universalism. *Television and New Media*, *20*(4), 319–335. https://doi.org/10.2139/ssrn.3384569

Mills, S. (2019). Who owns the future? Data trusts, data commons, and the future of data ownership. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3437936

Montgomery, J. (2017). Data sharing and the idea of ownership. *The New Bioethics*, *23*(1), 81–86. https://doi.org/10.1080/20502877.2017.1314893.

Open Data Institute. (2018, March 24). *No one owns data: We need to strengthen our rights*. https://theodi.org/article/no-one-owns-data-we-need-to-strengthen-our-rights/

Open Data Institute. (2019, June 13). *For stronger data rights, we must start with a shared language of data*. https://theodi.org/article/for-stronger-data-rights-we-must-start-with-a-shared-language-of-data/

Ostrom, E. (1990). *Governing the commons*. Cambridge University Press.

Ostrom, E. (2007). A diagnostic approach for going beyond panaceas. *Proceedings of the National Academy of Sciences*, *104*(39), 15181–15187. https://doi.org/10.1073/pnas.0702288104

Ostrom, V., Tiebout, C. M., & Warren, R. (1961). The organization of government in metropolitan areas: A theoretical inquiry. *American Political Science Review*, *55*(4), 831–842. https://doi.org/10.1017/s0003055400125973

P2PFoundation. (2019). *Abundance vs. scarcity*. https://wiki.p2pfoundation.net/Abundance_vs._Scarcity#Discussion_1:_Objective_Aspects

Pasquale, F. (2014). *IP law book review: Configuring the networked self: Law, code, and the play of every day practice. DigitalCommons@UM Carey Law*, *4*(1), 32–46. https://core.ac.uk/download/pdf/56358736.pdf

Peters, M. (2020). Digital socialism or knowledge capitalism? *Educational Philosophy and Theory*, *52*(1), 1–10. https://doi.org/10.1080/00131857.2019.1593033

Pierson, C. (2013). *Just property: A history in the Latin West*. Oxford University Press.

Polanyi, K. (1944). *The great transformation*. Rinehart and Co.

Prainsack, B. (2019a). Logged out: Ownership, exclusion and public value in the digital data and information commons. *Big Data & Society*, *6*(1). https://doi.org/10.1177/2053951719829773

Prainsack, B. (2019b). Data donation: How to resist the iLeviathan. In J. Krutzinna & L. Floridi (Eds.), *The ethics of medical data donation*. Philosophical studies series, Vol. 137 (pp. 9–22). Springer.

Prewitt, M. (2021). *A view of the future of our data*. www.noemamag.com/a-view-of-the-future-of-our-data/

Purtova, N. (2011). *Property rights in personal data: A European perspective*. BOXPress BV.

Purtova, N. (2017). Do property rights in personal data make sense after the big data turn: Individual control and transparency. *Journal of Law and Economic Regulation*, *10*(2), 64–78. https://pure.uvt.nl/ws/portalfiles/portal/26263149/01_The_Netherlands_Purtova_proofs.pdf

Ringstrom, A., & Vinocur, N. (2009, October 12). *A Nobel first: Economics prize goes to a woman*. Reuters. www.reuters.com/article/idINIndia-43100220091012

Roessler, B., & Mokrosinska, D. (2013). Privacy and social interaction. *Philosophy & Social Criticism*, *39*(8), 771–791. https://doi.org/10.1177/0191453713494968

Schermer, B. W., Custers, B., & van der Hof, S. (2014). The crisis of consent: How stronger legal protection may lead to weaker consent in data protection. *Ethics and Information Technology* *16*(2), 171–182. https://doi.org/10.1007/s10676-014-9343-8

Schwartz, P. (2019). Global data privacy: The EU way. *New York University Law Review*, *94*(4), 771–818.

Šestáková, A., & Plichtová, J. (2019). Contemporary commons: Sharing and managing common-pool resources in the 21st century. *Human Affairs*, *29*(1), 74–86. https://doi.org/10.1515/humaff-2019-0007

Smichowski, B. (2016). *Data as a common in the sharing economy: A general policy proposal*. CEPN – Centre d'Economie de l'Université Paris Nord. https://hal.archives-ouvertes.fr/hal-01386644

Smichowski, B. (2018). *The value of data: An analysis of closed-urban-data-based and open data-based business models*. Working Paper No. 1, SciencesPo. https://hal.archives-ouvertes.fr/hal-01736484/document

Smichowski, B. (2019). Alternative data governance models: Moving beyond one-size-fits-all solutions. *Intereconomics*, *54*(4), 222–227. http://dx.doi.org/10.1007/s10272-019-0828-x

Smith, H., & Burrows, R. (2021). Software, sovereignty and the post-neoliberal politics of exit. *Theory, Culture & Society*, *38*(6), 143–166. https://doi.org/10.1177/0263276421999439

Stavrides, S., & De Angelis, M. (2010, June). On the commons: A public interview with Massimo De Angelis and Stavros Stavrides. *e-flux, Journal*, #17. www.e-flux.com/journal/17/67351/on-the-commons-a-public-interview-with-massimo-de-angelis-and-stavros-stavrides/

Taylor, L., Floridi, L., & Sloot, B. (2017). *Group privacy: New challenges of data technologies*. Springer.

Taylor, M. (2006). Data protection: Too personal to protect? *SCRIPT-Ed*, 71–81. https://doi.org/10.2966/scrip.030106.71

The Data Freedom Act. (2020). www.radicalxchange.org/wp-content/uploads/2020/03/DFA.pdf

Tisne, M. (2018). It's time for a Bill of Data Rights. *MIT Technology Review*, 25. www.technologyreview.com/2018/12/14/138615/its-time-for-a-bill-of-data-rights/

Ubaldi, B., Van Ooijen, C., & Welby, B. (2019). *A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance*. Organisation for Economic Cooperation and Development (OECD).

Urban, M. (2019). *New rules for the game: Rebooting Canada's competition regime for the digital economy*. Mowat Centre.

Vaidhyanathan, S. (2003). *Copyrights and copywrongs: The rise of intellectual property and how it threatens creativity*. New York University Press.

Waldron, J. (1988). *The right to private property*. Clarendon Press.

Yakowitz, J. (2011). Tragedy of the data commons. *Harvard Journal of Law and Technology*, *25*, 1–67. https://doi.org/10.2139/ssrn.1789749

Zhang, G.-Q., Cui, L., Mueller, R., Tao, S., Kim, M., Rueschman, M., & Redline, S. (2018). The National Sleep Research Resource: Towards a sleep data commons. *Journal of the American Medical Informatics Association*, *25*(10), 1351–1358. https://doi.org/10.1093/jamia/ocy064

Zuboff, S. (2019). *Age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.

Zumbansen, P. (2010). Transnational legal pluralism. *Transnational Legal Theory*, *10*(2), 141–189. https://doi.org/10.1080/20414005.2010.11424506