Bottom-up data Trusts: disturbing the 'one size fits all' approach to data governance

Sylvie Delacroix * and Neil D. Lawrence**

Key Points

- The current lack of legal mechanisms that may plausibly empower us, data subjects to 'take the reins' of our personal data leaves us vulnerable. Recent regulatory endeavours to curb contractual freedom acknowledge this vulnerability but cannot, by themselves, remedy it—nor can data ownership. The latter is both unlikely and inadequate as an answer to the problems at stake.
- We argue that the power that stems from aggregated data should be returned to individuals through the legal mechanism of Trusts.
- Bound by a fiduciary obligation of undivided loyalty, the data trustees would exercise the data rights conferred by the GDPR (or other top-down regulation) on behalf of the Trust's beneficiaries. The data trustees would hence be placed in a position where they can negotiate data use in conformity with the Trust's terms, thus introducing an independent intermediary between data subjects and data collectors.
- Unlike the current 'one size fits all' approach to data governance, there should be a plurality of Trusts, allowing data subjects to choose a Trust that reflects their aspirations, and to switch Trusts when needed. Data Trusts may arise out of publicly or privately funded initiatives.
- By potentially facilitating access to 'pre-authorized', aggregated data (consent would be negotiated on a collective basis), our data Trust proposal may remove key obstacles to the realization of the potential underlying large datasets.

Introduction

From the friends we make to the foods we like, via our shopping and sleeping habits, most aspects of our quotidian lives can now be turned into machine-readable data points. For those able to turn these data points into models predicting what we will do next, this data can be a source of wealth. For those keen to replace biased, fickle human decisions, this data—sometimes misleadingly—offers the promise of automated, increased accuracy. For those intent on modifying our behaviour, this data can help build a puppeteer's strings. As we move from one way of framing data governance challenges to another, salient answers change accordingly. Just like the wealth redistribution way of framing those challenges tends to be met with a property-based, 'it's *our* data' answer, when one frames the problem in terms of manipulation potential, dignity-based, human rights answers rightly prevail (via fairness and transparency-based answers to contestability concerns). Positive data-sharing aspirations tend to be raised within altogether different conversations from those aimed at addressing the above concerns. Our data Trusts proposal challenges these boundaries.

This article proceeds from an analysis of the very particular type of vulnerability concomitant with our 'leaking' data on a daily basis, to show that data ownership is both unlikely and inadequate as an answer to the problems at stake. We also argue that the current construction of top-down regulatory constraints on contractual freedom is both necessary and insufficient. To address the particular type of vulnerability at stake, bottom-up empowerment structures are needed. The latter aim to 'give a voice' to data subjects whose choices when it comes to data governance are often reduced to binary, ill-informed consent. While the rights

* Sylvie Delacroix, Alan Turing Institute and University of Birmingham (Law).

** University of Cambridge. We are grateful to Michael Veale, Gianclaudio Malgieri, Ben McFarlane, and Robert Chambers for their insightful comments and suggestions.

granted by instruments like the GDPR can be used as tools in a bid to shape possible data-reliant futures—such as better use of natural resources, medical care, etc, their exercise is both demanding and unlikely to be as impactful when leveraged individually. As a bottom-up governance structure that is uniquely capable of taking into account the vulnerabilities outlined in the first section, we highlight the constructive potential inherent in data Trusts. This potential crosses the traditional boundaries between individualist protection concerns on one hand and collective empowerment aspirations on the other.

The second section explains how the Trust structure allows data subjects to choose to pool the rights they have over their personal data within the legal framework of a data Trust. It is important that there be a variety of data Trusts, arising out of a mix of publicly and privately funded initiatives. Each Trust will encapsulate a particular set of aspirations, reflected in the terms of the Trust. Bound by a fiduciary obligation of undivided loyalty, data trustees will exercise the data rights held under the Trust according to its particular terms. In contrast to a recently commissioned report,[1] we explain why data can indeed be held in a Trust, and why the extent to which certain kinds of data may be said to give rise to property rights is neither here nor there as far as our proposal is concerned. What matters, instead, is the extent to which regulatory instruments such as the GDPR confer rights, and for what kind of data. The breadth of those rights will determine the possible scope of data Trusts in various jurisdictions.

Our 'Case Studies' aim to illustrate the complementarity of our data Trusts proposal with the legal provisions pertaining to different kinds of personal data, from medical, genetic, financial, and loyalty card data to social media feeds. The final section critically considers a variety of implementation challenges, which range from Trust Law's cross-jurisdictional aspects to uptake and exit procedures, including issues related to data of shared provenance. We conclude by highlighting the way in which an ecosystem of data Trusts addresses ethical, legal, and political needs that are complementary to those within the reach of regulatory interventions such as the GDPR.

From big mother to big brother and vice-versa: risk or asset?

The promise of the modern digital society is that our computers will be able to second guess us, providing for our every need like some form of digital mother. Such omniscient provision[2] presupposes a degree of surveillance that can quickly flip from well-intentioned benevolence to the malign curbing of individual freedoms—a form of Big Brother—depending on the way you look at it: just like two sides of the same coin (Big Mother v Big Brother). The extent to which we are each comfortable with such digital supervision is inherently subjective, and therefore surely a matter for personal choice, but our levers for control are currently limited: at times we can turn the lever to on or off. At others, the lever seems out of reach entirely.

Another way of formulating the 'Big Brother/Big Mother' dichotomy relies on the concept of risk: from a Big Brother perspective, before becoming an asset, personal data is first and foremost a source of risk. A risk that is unfamiliar to us, and difficult to assimilate in our decision making. The unfamiliarity of this risk is not restricted to the data subjects: as a law-making body that may be tasked with reforming data governance, the American Senate's hearing after the Cambridge Analytica scandal was almost as newsworthy as the scandal itself, revealing a very poor understanding of Facebook's business model.[3]

Instrumental-risks in this data proliferation have been understood for a number of years. Arthur Miller's 'The Assault on Privacy' was written in the late 1960s and outlines many of the risks that data protection legislation attempts to ameliorate. In the intervening years, a different kind of risk, which is cumulative by nature, has emerged: the data we leak daily has become something by reference to which we may be continuously judged.[4] The systematic collection of data allows our lives to be dissected to an unprecedented degree. Although any individual fact learned about us may be

1 Chris Reed, BPE solicitors and Pinsent Masons, *Data trusts: Legal and Governance Considerations* (2019). <https://theodi.org/wp-content/uploads/2019/04/General-legal-report-on-data-trust.pdf> accessed 10 September 2019.

2 V Mavroudis and M Veale, 'Eavesdropping whilst you are Shopping: Balancing Personalisation and Privacy in Connected Retail Spaces' (2018) 18 IET Conference Proceedings 10.

3 Sara Fischer and Dan Primack, 'Mark Zuckerberg outwits Congress' (*AXIOS*, 2018) <https://www.axios.com/mark-zuckerberg-outwits-congress-facebook-42fc1d21-ba2f-4cbb-82a2-93c29bae969c.html> accessed 10 September 2019.

4 Hildebrandt advocates a practice of 'agonistic machine learning' to provide us with the 'means to achieve effective protection against overdetermination of individuals by machine inferences': Mireille Hidebrandt, 'Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning' (2019) 20 Theoretical Inquiries in Law 83.

inconsequential, taken together, over time, a detailed picture of who we are and what motivates us emerges.[5]

It is not a new phenomenon for decisions to be taken about us on the basis of our personal data. Some of these decisions can be challenged by legal means. Yet, the amount of time and effort required to mount such challenges means they are likely to be restricted to decisions of considerable import. While the European situation[6] may not be as bad as elsewhere[7] when it comes to accessibility issues, the active exercise of one's data protection rights nevertheless requires a considerable level of knowledge and agility. Independently of these accessibility concerns, ex-post remedies remain poorly suited to a world where a vast number of seemingly insignificant decisions can together end up painting a radically different picture for our lives.

The cumulative way in which we provide our data, as well as the cumulative nature of the resulting decision-making, presents particular challenges to our freedoms.[8] Never before has the self we aspire to be been constrained to such an extent by our past[9]: not just in instrumental ways, as recognized by the right to be forgotten, but in a subtly nefarious manner. It is the insidious nature of this risk that is not well addressed[10] by our existing legal mechanisms. How do we control for this death by a thousand cuts?

Data ownership?

One of the mechanisms through which we have traditionally sought to assert control over our surroundings is *ownership*. As an instrument of control, the concept of ownership has facilitated past aberrations such as the ability to own other people (slavery) or to restrict the extent to which people may enjoy the fruits of their labour (serfdom). Ownership has also been taken as a proxy for other forms of power, as instantiated in the restriction of voting rights to land-owners. Today *data ownership* is sometimes hailed as a precondition in order to return 'control' to the individual (particularly so in the American literature[11]). This ongoing, intuitive association of ownership with control seems to draw on a specific ideal of property, which is reflected in the saying, 'one's home is one's castle': 'in the usual course of events, access to a person's home requires a consensual transaction with the owner, and unconsented uses can be enjoined'.[12]

Yet, as Evans points out, 'different assets call for different forms of ownership' (or more accurately, different types of property rights). Personal data is rather unlike homes (or castles). The type of property rights data can give rise to[13] are more akin to the 'nonexclusive rights riparian owners have in a river that runs by their land'.[14] Not only can public good considerations justify substantial restrictions on the use of that stretch of river (just like they can for data[15]). The necessary inter-dependence between those located upstream and downstream entails the need to take into account others' right to non-interference. This riparian metaphor is helpful when considering a variety of regimes governing conflicting data rights, such as that pertaining to medical data used for research purposes (navigation rights trumping irrigation rights could be compared to researchers' right to decline an erasure request).

Ownership is not only unlikely to provide the level of control[16] wished for: it is also a poor answer to the

5 Mireille Hildebrandt, 'Defining Profiling: A New Type of Knowledge?' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen: Cross-Disciplinary Perspectives* (Springer, Netherlands 2008).

6 In Europe, individuals can lodge data protection complaints with their national supervisory authority—the ICO in the UK. The GDPR's making it possible for Member States to implement collective redress mechanisms (art 80) also has the potential to improve accessibility issues.

7 Under the CCPA in the USA, class actions would currently be limited to data breaches only (though there are calls for the private right of action to be broadened beyond such data breaches): the California Consumer Privacy Act 2018 (if it comes into force without amendments in January 2020) would mean that the need for the plaintiff to establish that she has suffered actual harm following a data breach would be bypassed, allowing for the filing of class actions to obtain statutory damages following a data breach.

8 Julie Cohen, *Configuring the Networked Self* (Yale University Press, New Haven and London 2012).

9 S Delacroix and M Veale, 'Smart Technologies and Our Sense of Self: The Limitations of Counter-Profiling' in Mireille Hildebrandt and Kieron O'Hara (eds), *Life and the Law in the Era of Data-Driven Agency* (Edward Elgar, Cheltenham 2020).

10 Wachter and Mittelstadt emphasize that 'compared to other types of personal data, inferences are effectively 'economy class' personal data in the GDPR': S Wachter and B Mittelstadt, 'A right to reasonable inferences: re-thinking data protection law in the age of Big Data and AI' (2019) 2 Columbia Business Law Review 494.

11 See also JB Rule and L Hunter, 'Towards Property Rights in Personal Data' in CJ Bennett and R Grant (eds), *Visions of Privacy: Policy Choices for the Digital Age* (University of Toronto 1999) and JB Rule, *Privacy in Peril* (OUP 2007) 196. For a European voice arguing along the same lines, see N Purtova, *Property Rights in Personal Data: A European Perspective* (Kluwer Law International 2012).

12 Barbara J Evans, 'Much Ado About Data Ownership' (2011) 25 Harvard Journal of Law and Technology 69.

13 This will be expanded upon in section 'The possibility (and advantages) of holding data rights under a legal Trust'.

14 Evans.

15 'Individual control and property rights over personal data are both on the same spectrum of potential regulatory responses to the personal data processing phenomenon [. . .] If control is absolute, it will diminish the public domain and have a negative impact on other rights and interests. Moreover, the realities of the bargaining process and of the technological environment cannot be ignored. This is something which must be borne in mind by those advocating strong rights of control or rights of ownership on behalf of individual data subjects, and when brining to the market personal data lockers and vaults': Orla Lynskey, *The Foundations of EU Data Protection Law* (OUP 2018) 252.

16 Along this line, Lazaro and Le Metayer emphasise that '[t]he premise of autonomy and active agency implied in this rhetoric [of control] seems to be radically undermined in the context of contemporary digital environments and practices'; Christophe Lazaro and Daniel Le Metayer,

type of problems (and vulnerabilities) at stake.[17] While some of the early voices that were instrumental to the emergence of data privacy law advocated extensive reliance on property law,[18] today the 'law and doctrine on human rights' are 'generally regarded as providing the principal normative basis'[19] for data privacy law. This turn to human rights reflects the fundamental nature of the harms that can ensue from the abusive exploitation of personal data. Not only are such harms ill-addressed[20] through material compensation; they are also difficult, if not impossible, to prevent through individual vigilance alone. Unlike most homes (or river) owners, very few of us have the time or know-how to understand—let alone control—what parts of our data we are happy to share, and under what terms. The systematic monitoring of one's data presupposes resources that most simply do not have.

Regulatory interventions such as the General Data Protection Regulation (GDPR),[21] which divides us primarily into data subjects and data controllers, seek to address these epistemic imbalances. The term data subject has unfortunate, but perhaps appropriate, connotations of royalty and feudal society, where an individual is subject to whims beyond their control: there is a power-asymmetry between the subjects and the controllers.[22] This asymmetry arises because the controllers have accumulated data from many individuals, which allows them to invest time and expertise into the processing. In contrast, the subject has knowingly or unknowingly provided her data to many entities—including mostly 'invisible' data brokers[23]—and has neither the expertise nor the time to unpick each data controller's motivations and methods.

The downsides of a contractual approach to data governance

The duties imposed by the GDPR on data controllers stem in part from an acknowledgement that individual data subjects are rarely in a position to bargain. Aside from the need to address gross power imbalances, those curbs on contractual freedom also proceed from an awareness of the fact that a strictly contractual approach[24] to data governance is likely to compromise aspirations that underlie the very *raison-d'être* of liberal democracies.

Among these aspirations is a commitment to the rejection of social cruelty.[25] Our digital society generates novel forms of vulnerabilities: today it is near impossible to go about one's life without leaving a data trail that, potentially, reveals more about ourselves to strangers than we'd ever disclose to friends. This makes us vulnerable. This vulnerability can be exploited in a way that compromises our ability to retain some minimal sense of 'authorship' over our lives: our ability to maintain a social self that is at least partly set out by us is compromised, just as the vulnerability of the elderly, the ill, or the prosecuted can be overlooked with similar effects. Their institutionally exposed age or weakness can be ignored—or taken advantage of—resulting in the compromission of 'their capacity to develop and maintain an integral sense of self'.[26] The loss of opacity that is concomitant with the potentially detailed, personal knowledge garnered by data controllers is at least as conducive to similar forms of social cruelty, albeit in less visible, more subtle ways.

It is because healthcare providers, lawyers, and educators are in a position to greatly exacerbate—or moderate—the above vulnerability that high standards of

'Control over personal data: true remedy or fairy tale?' (2015) 12 SCRIPTed 3, 29.

17 Kan and Buchner also highlight the extent to which the adoption of a 'property rights' terminology to conceptualize our personal data can have insidious effects, potentially leading us to treat our personal data more 'like [our] car than [our] soul' Jerry Kan and Benedikt Buchner, 'Privacy in Altantis' (2004) 18 Harvard Journal of Law and Technology 229, 260.

18 AF Westin, *Privacy and Freedom* (Atheneum 1967).

19 LA Bygrave, *Data Privacy Law* (OUP 2014).

20 The explicit inclusion of 'non-material' damages within the scope of eligible damages in art 82 GDPR has been welcomed as providing much needed clarity.

21 EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 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), OJ 2016 L 119/1. Due to space constraints, we have not addressed in any detail the way in which our proposed data trusts would complement other regulatory interventions, such as the California Consumer Privacy Act 2018 (CCPA): building upon distinct frameworks of data protection developed since the 1970s, the CCPA endows 'consumers' with rights against 'businesses' holding their 'personal data' (which is broadly defined).

22 Recital 43 refers to the 'imbalance' between data controller and data subject. See also Paul De Hert and Serge Gutwirth, 'Privacy, data protection and law enforcement. Opacity of the individual and transparency of power' in E Claes, A Duff and Serge Gutwirth (eds), *Privacy and the Criminal Law* (Intersentia 2006).

23 Data brokers are not directly susceptible to consumer pressure in the manner of consumer-facing data controllers. Sociologists have criticised the lack of regulation in the industry: Leanne Roderick, 'Discipline and Power in the Digital Age: The Case of the US Consumer Data Broker Industry' (2014) 40 Critical Sociology 729.

24 Volokh has argued that we can protect privacy by private ordering through contract: Eugene Volokh, 'Freedom of Speech and Information Privacy: The Troubling Implications of a Right to Stop People from Speaking About You' (2000) 52 Stanford Law Review 1049.

25 Sangiovanni defines social cruelty as involving 'the unauthorized, harmful, and wrongful use of another's vulnerability to attack or obliterate their capacity to develop and maintain an integral sense of self': Andrea Sangiovanni, *Humanity without Dignity: Moral Equality, Respect and Human Rights* (Harvard UP 2017). Note that the term 'use' is to be understood loosely: Sangiovanni for instance refers to instances of institutional neglect (such as in residential care homes).

26 Ibid.

professional responsibility are expected of them. What about those in charge of deciding what parts of our personal data can be collected and processed, and for what purpose? They too are in a position to greatly exacerbate the particular vulnerability[27] that flows from the disclosure of granular information about our machine-readable past (and present). Yet, unlike healthcare providers, lawyers, or educators, this position of power is not concomitant with any personal relationship, making it all the easier for data controllers to insulate themselves from the responsibility underlying their position. Data protection law foresees many situations in which an organization will hold data about an individual without any direct contact having ever been made.[28]

Summary so far

We have characterized some of the challenges associated with the cumulative and insidious way in which we provide our data:

1. Recent regulatory endeavours to curb contractual freedom cannot by themselves reverse the power-asymmetry between data controllers/businesses and data subjects/consumers.
2. The current lack of legal mechanisms that may plausibly (and collectively) empower data subjects threatens one of the key commitments underlying our liberal democracies. The vulnerability that stems from the institutional exposure of our machine-readable past can be overlooked or exploited in a way that compromises our ability to maintain a *social self* that is at least partly controlled by us. Unlike the instrumental risks attached to punctual decisions (such as a mortgage applications), the risks concomitant with this vulnerability are cumulative: seemingly inconsequential decisions have downstream effects outside our control.
3. The above risks are also difficult to grasp and inherently subjective. The current, 'one size fits all' approach does not allow different individuals to choose among different approaches to data governance, which reflect both their subjective attitude to risk and their moral and political aspirations.

In the next section, we consider how we can address these challenges through the mechanism of data Trusts.

The Trust structure as a way of taking into account the vulnerabilities at stake

We propose data Trusts as a bottom-up mechanism, whereby data subjects choose to pool the rights they have over their personal data within the legal framework of the Trust. In our proposal, the data subjects tend to be both the *settlors* and the *beneficiaries* of the Trust: the trustees are compelled to manage the subjects' data according to the terms of the Trust. They have a *fiduciary* responsibility towards the data subjects (the beneficiaries of the Trust). We envision an ecosystem of Trusts arising out of a mix of publicly and privately funded initiatives,[29] each with different constitutional terms, allowing data subjects to choose among different approaches to data governance. A successful Trust would be in control of more data and be able to deliver more benefit to data subjects.

To the best of our knowledge, the idea of 'bottom-up' data Trusts was first publicly suggested in 2016.[30] The legal mechanism of a data Trust[31] aims to leverage the resources concomitant with the pooling of data to directly address the power-asymmetries mentioned above. A Trust is formed when a person in whom a set of resources is vested—the Trustee—is compelled to hold and manage those resources either for the benefit of another person(s)—the beneficiaries—or for some legally enforceable purpose(s) *other than the Trustee's own*. Aside from its allowing for 'more subtle shades of ownership than the common law permits',[32] the duties which a Trust structure imposes upon the Trustee(s) are also better suited to the particular vulnerabilities at stake, as they demand the Trustee's undivided loyalty and dedication to the interests and aspirations of the data subjects (as beneficiaries of the Trust).

The Trustee's duties are *fiduciary*. A fiduciary duty is considerably more onerous than a 'duty of care'. Under Tort law, a defendant may be found in breach of that duty if the defendant is shown to not have taken

27 Coeckelbergh examines the extent to which ICT transforms the way in which we cope with 'existential vulnerabilities'. This Heideggerian understanding of vulnerability differs from the notion of vulnerability explored here: Mark Coeckelbergh, 'The Art of Living with ICTs: The Ethics–Aesthetics of Vulnerability Coping and Its Implications for Understanding and Evaluating ICT Cultures' (2017) 22 Foundations of Science 339.

28 Art 14(5)(b) of the GDPR allows for controllers to never contact an individual when they indirectly receive data relating to them insofar as it involves 'disproportionate effort'.

29 Implementation challenges, including the successful 'seeding' of an ecosystem of data Trusts, are discussed in section 'Implementation challenges'.

30 Neil Lawrence, 'Data trusts could allay our privacy fears', *The Guardian* Media & Tech Network <https://www.theguardian.com/media-network/2016/jun/03/data-trusts-privacy-fears-feudalism-democracy> accessed 10 September 2019.

31 In 'Implementation challenges', we explain why the legal mechanism of a data Trust can be relied on, in part addressing some of the questions raised in Kieron O'Hara, 'Data Trusts: Ethics, Architecture and Governance for Trustworthy Data Stewardship' (2019 Web Science Institute White Papers).

32 Scott Atkins, *Equity and Trusts* (OUP 2018).

'reasonable care' to avert some 'reasonably foreseeable' harm. Reasonable care is defined as the level of concern that a 'reasonable person'[33] would apply. Our contention is that, where personal data is concerned, 'reasonable care' is not sufficient: the 'ethical relationship of trust'[34] underlying fiduciary duties is better suited to the vulnerabilities at stake, especially given the Court's jurisdiction to supervise and, if necessary, to intervene on behalf of the beneficiaries to enforce the Trust against a trustee. The fact that the burden of proof is reversed is also important. In the event of a claim, it is for the trustees to demonstrate that they have sought to promote the beneficiaries' interest with appropriate degrees of impartiality, prudence, transparency and undivided loyalty.[35] This loyalty condition is important: while trustees can be remunerated for their services, trustees cannot profit from a Trust, and more generally cannot allow their own interests to conflict with those of the beneficiaries. This condition calls into question those proposals that seek to impose fiduciary duties upon data collectors/controllers.

Among those proposals, Edwards was the first to see the potential inherent in Trusts as a legal mechanism for managing a resource—personal data—that is accumulated by small contributions.[36] Under her proposal, an 'implied'[37] data Trust is created whenever data subjects share personal data with data collectors. The latter are deemed to be the trustees. This is problematic, given the undivided loyalty[38] condition. In contrast, our proposal presupposes the creation of 'express' Trusts, and it requires the appointment of independent trustees who are bound by the Trust's purposes and terms. The latter can vary from one Trust to another. Importantly, this allows for an ecosystem of Trusts, where a variety of data sharing policies across Trusts gives data subjects a range of choices that reflect their personal trade-offs: the resulting diversity also allows society to explore different principles for data sharing within the same digital ecosystem.

Information fiduciaries

Balkin suggests that economic and tax incentives[39] ought to be offered to data controllers in exchange for their accepting 'fiduciary obligations', provided these obligations are 'not too broad'. Why not too broad? Because 'it might follow that online service providers could not make any money at all from this data because the data might be used in some way to some end-user's disadvantage'.[40]

This proposal does not tackle the power asymmetries inherent in our current system of data feudalism. Balkin is right to point out that 'fiduciary' does not mean 'not for profit': in many jurisdictions[41] healthcare providers and lawyers are deemed to have fiduciary obligations towards their patients or clients. This does not mean that they cannot be paid for their work. What it does mean is that they have an obligation of *undivided loyalty* towards their patient/client: a doctor who is set to profit from her patients' drug prescriptions (because of her holding pharma shares, say) would be found in breach of her fiduciary duties.[42]

If a data controller has a business interest in the data provided by data subjects, this results in a conflict between that interest and her duty towards data subjects. Data controllers in this position would be obliged to both maximize the value of the personal data they collect (for the benefit of shareholders) *and* concomitantly honour fiduciary obligations towards data subjects. While Balkin does acknowledge the potential for conflict of interest,[43] he fails to draw the only logical conclusion: a fiduciary obligation towards data subjects is incompatible with the data controllers' responsibility towards shareholders. As discussed above, to honour a fiduciary obligation not only demands independence from profit maximization: it also requires an ability to relate to the complex and multi-faceted nature of the

33 John Gardner, 'The many Faces of the Reasonable Person' (2015) 131 Law Quarterly Review 563.

34 Hartley Goldstone, 'The Moral Core of Trusteeship' (2013) 152 Trusts & Estates 5.

35 The delineation of a Trustee's duties and responsibilities is open to interpretation. It varies to some extent from one jurisdiction to another, but also in accordance with the terms and purpose of the Trust. Most importantly, implementation choices (discussed in section 'Challenges and opportunities inherent in holding data rights under a Trust (rather than a contractual agreement or corporate structure') will determine whether the trustee must be deemed a data controller under the GDPR or not.

36 L Edwards, 'The Problem with privacy' (2004) 18 International Review of Law Computers & Technology.

37 Given Edwards' referring to the 'administrative headache' that would stem from the fact that 'millions of data subjects would have a claim on millions of different aggregate data 'trusts', with one trust for each data collector', it is unlikely that such Trusts would be express Trusts. See ibid.

38 There is a degree of confusion about the exact nature of the 'trustee's' duties under Edwards' scheme: under Equity law, the latter cannot but be fiduciary, with the stringent obligations discussed above. It may be that the ambiguity stems in part from the fact the Trust structure is referred to by Edwards as a way of justifying a 'privacy tax', which in practice does not require any reliance on Equity (in effect, the Trust structure mostly serves a justificatory purpose).

39 Jack M Balkin, 'Information Fiduciaries and the First Amendment' (2016) 49 UC Davis Law Review 1183. Zittrain also suggests immunity from certain kinds of lawsuits among the incentives that could be offered: Jonathan Zittrain, 'Engineering an Election' (2013) 127 Harvard Law Review Forum 335, 339.

40 Balkin (n 39) 1227.

41 Canada is the only country to explicitly characterize the patient–doctor relationship as fiduciary. In the US fiduciary duties are recognized as an independent ground for action in respect to particular obligations (such as confidentiality) in some states.

42 *Moore v Regents of the Univ of Cal*, 793 P.2d 479 (Cal 1990).

43 Balkin (n 39) 1126.

vulnerability inherent in the data subject/data controller relationship. In this respect, the 'information fiduciary' proposed by Balkin would be placed in a position that is comparable to that of a doctor[44] who gains a commission on particular drug prescriptions or a lawyer who uses a company to provide medical reports for his clients while owning shares in that company.[45]

Aside from sidestepping the conflict of interest issue mentioned above, Balkin's information fiduciary proposal only affords protection to those who are already in a contractual relationship with 'digital companies'. Balkin acknowledges this issue in his more recent paper, noting that: 'there are a wide range of situations in which people lack a contractual relationship with a digital enterprise or with a business that collects personal information and uses algorithms to make decisions.'[46] To address this issue, Balkin puts forward the Common Law concepts of public and private nuisance: in Balkin's argument, the nuisance that can legitimately be targeted by regulation (despite First Amendment constraints) consists in '[u]sing algorithms repeatedly and pervasively over large populations of people [which] may inappropriately treat people as risky or otherwise undesirable, impose unjustified burdens and hardships on populations, and reinforce existing inequalities'.[47] If such is the nature of the 'nuisance' at stake, the remedy proposed by Balkin is puzzling:

> The appropriate remedy is to make companies internalize the costs they shift onto others and onto society as a whole as they employ algorithmic decision making.[48]

Balkin does not dwell on the process that would somehow enable the quantification (and hence 'internalization') of the 'cost' of treating people as 'otherwise undesirable'. The type of inequality that is fostered by the current, 'feudal' approach to data governance is not merely one of material resources or opportunities. As discussed above, the compromising of our ability to maintain a *social self* that is at least partly controlled by us undermines our commitment to *moral* equality: our equal moral worth independently of any contingent traits (such as the amount of personal data shared online). Taxes or economic incentives can sometimes prove effective when tackling socio-economic inequalities. They are a dubious answer when faced with structures that foster what may aptly be described as a form of social cruelty.

Instead of seeking to compensate for the nefarious effects of the existing data governance framework, our proposal is to challenge it from the ground-up. As a bottom-up structure, the data Trust framework is in a position to plausibly empower data subjects, to 'take the reins' of their data. This contrasts with the current focus on compensation for the undesirable risks or side-effects that stem from the current exploitation of our data by centralized platforms.

When 'Trust' is used as a marketing tool

The term 'data trust' has also been used recently to refer to the need for some 'repeatable framework of terms and mechanisms', 'to facilitate the sharing of data between organizations holding data and organizations looking to use data to develop AI'.[49] Such frameworks have an important role to play in facilitating the responsible sharing of data, which may otherwise remain out of the reach of organizations that do not have the requisite type of know-how (and legal support) to be able to leverage such data without fear of breaching ethical and legal requirements. But it is unclear what, if anything, such frameworks have in common with legal Trust structures.

In contrast to the above proposals, the 'data Trusts' we have in mind are 'true' Trusts, legally speaking.[50] The collective setting of terms by the Trust is a way for data subjects to pool their rights to acquire a 'voice': some Trusts may indeed be run in a way that resembles a collective or cooperative. There are many historical precedents for the formation of such bodies to empower the disenfranchised. For example, 'Land Societies'[51] were formed almost two centuries ago for the purpose of giving a political voice to their members, who would not otherwise have had the resources to acquire the freehold land conditioning their right to vote.

The terms of the Trust may specify a governance structure that compels the data Trustees to continuously consult and deliberate with the settlors and beneficiaries: in that case the Trust model would effectively function in ways similar to a cooperative, albeit with robust fiduciary

44 Balkin's proposal has the merit of acknowledging some of the similarities between the vulnerability that characterises the doctor/patient (or lawyer/client) relationship and that which underlies the data subject/data controller relationship, even if Balkin only focuses on the epistemic aspect of data subjects' vulnerability: ibid.

45 *Solicitors Regulation Authority v Dennison* [2012] EWCA Civ 421.

46 Jack M Balkin, 'Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation Essays' (2017) 51 UC Davis Law Review 1149.

47 Ibid.

48 Ibid.

49 Wendy Hall and Jérôme Pesenti, *Growing the Artificial Intelligence Industry in the UK* (2017).

50 For detailed arguments as to why data Trusts can be 'true' Trusts, see section 'The possibility (and advantages) of holding data rights under a legal Trust'.

51 Thomas Beggs, 'Freehold Land Societies' [[Royal Statistical Society, Wiley]] 16 Journal of the Statistical Society of London 338.

duties vested in the Trustees. Some data Trusts may prefer a less participatory model, while others may choose to automatically align themselves to a set of default data governance principles set by some governmental body.

This 'bottom-up' data Trust model is resolutely complementary to top-down, regulatory constraints (including those of the GDPR). Indeed, some of those top-down constraints will be needed in relation to public good issues: one may for instance consider an across-Trusts constraint which demands that some kind of health-data be shared in all cases. The choices and aspirations encapsulated in each data Trust will necessarily be limited by top-down, public interest interventions which must delineate the scope of legitimate discretionary choice. In this respect, data trustees are likely to play an important role in shaping societal debate about 'whether one can require an individual [or group of individuals] to contribute to a "greater" good'.[52] We firmly believe in a positive answer to the latter question, which needs to be the focus of greater public awareness and debate than is currently the case.

Data trustees will also have to make sure that any data—with its relevant rights—entrusted by settlors is in fact genuinely theirs to give (and not merely 'captured' by a data collecting artefact deployed by the purported settlors, for instance). The latter responsibility could be formalized as part and parcel of a larger set of 'professional' duties (in which case data trustees would be overseen by a specific professional body), or alternatively as a duty that has to be complied with for the purpose of certification.

The need to be able to 'shop around' data Trusts

While we are unlikely to have a data Trust tailored specifically to each individual in society it seems important that there be a wide variety of data Trusts, each instantiating one particular way of balancing data risks and responsibilities. Some Trusts may heavily favour the furthering of some 'public good' endeavour by making some data freely accessible to some organizations, while others may prioritize the maximization of financial returns. Others may put great emphasis on minimizing individual risks. Individuals will be able to shop around, switching from one Trust to another as and when their preferences or aspirations evolve. The fostering of such competition between a wide variety of data Trusts will not only serve to raise awareness of the fact that there are many ways of apprehending data risks and responsibilities. It will also make it more likely that our data governance structures remain in touch with the evolving needs and aspirations of multi-faceted societies.

Two conditions must be met for such an ecosystem to thrive: the barrier for entry must be low—the creation of new Trusts must be relatively straightforward—(condition 1) and the data subjects' data must be secure (condition 2). Given that the expertise for building secure data infrastructure is in short supply, many Trusts may prefer to focus on collectively setting the terms according to which the settlors' data may be used, relying on computational and storage infrastructure from commercial suppliers. It might even be the case that data subjects retain physical control of their data, such as in a personal data container, but give Trustees the ability to exercise their rights and undertake operations over it.

The above also implies a system of data exchange between Trusts and consumers of the data (companies, hospitals, etc). Such a data exchange system requires two fundamental characteristics.

1. An individual's personal data must be portable between different computer systems (data portability—condition 3)
2. An individual's data must be erasable from any particular system (data erasure—condition 4)

The successful development of an ecosystem of Trusts is contingent on their ability to make use of the currently limited rights around data portability and data erasure: for any Trust to have power it must be able to make its settlors' data available (under its constitutional terms), but it must also be able to withdraw data from a particular controller if it is to specify disapproval of a particular form of data use.

At the moment, this need to be able to demand data erasure is only partially backed by legal provisions. In Europe, Article 17 of the GDPR grants a right to have personal data erased in certain circumstances only. Among the several limitations, the 'overriding legitimate interest to continue processing'[53] constitutes a significant exception.[54]

52 Hielke Hijmans and Charles Raab, 'Ethical dimensions of the GDPR', 20 August 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3222677, forthcoming in Mark Cole and Franziska Boehm (eds), Commentary on the General Data Protection Regulation (Edward Elgar).

53 When 'legitimate interest' is relied on as the basis of processing.

54 Similar exceptions apply to the right to deletion under the CCPA, which also includes an exemption from deletion requests if such requests interfere with a right to 'exercise free speech, ensure the right of another consumer to exercise his or her right of free speech, or exercise another right provided by law' Peter Stockburger, 'The Good, Bad, And The Ugly: Key Takeaways From California's New Privacy Law' (*Mondaq*, 13 November 2018) <http://www.mondaq.com/canada/x/754510/The+Good+Bad+And+The+Ugly+Key+Takeaways+From+Californias+New+Privacy+Law> accessed 10 September 2019. For a discussion of the trade-off between confidentiality requirements and the data controller's ability to enact the data subjects' rights to erase (or access) their data, see Michael Veale, Reuben Binns and Jef Ausloos, 'When Data Protection by Design and Data Subject Rights Clash' (2018) 8 International Data Privacy Law 105.

If the erasure requirement is a form of negative control, the portability requirement allows for positive reinforcement: data can be shared with actors that conform to the policies of the Trust. Under Article 20 of the GDPR, this portability right is limited to information that has been 'provided' to a controller, actively or passively, and processed on the basis of consent (or performance of a contract).[55] In contrast, the right of access (Article 15 of the GDPR) is stronger and covers all data, but this data can be provided in any format (unlike Article 20[56]), which might hinder automatic transfer or parsing.[57] Interestingly, article 16(4) of the very recent directive 'on certain aspects concerning contracts for the supply of digital content and digital services' broadens the scope of portability rights to non-personal data, as consumers are given the right to retrieve 'any content other than personal data, which was provided or created by the consumer when using the digital content or digital service supplied by the trader'.[58]

Now, the impact of those—currently limited—portability and access rights will in part depend upon shortening the timeframe within which they have to be acted upon. Article 15 GDPR specifies that a subject access request should be fulfilled within a month. For efficient data exchange we might expect Trusts to require access within milliseconds: performed programmatically,[59] such access requests would be executed by direct computer interface, without human involvement. Similar stipulations apply to data erasure.

The impact of those rights will also depend on quality control and compliance mechanisms: a recent study analysing the replies to a number of access requests highlights reports of a large variation in the quality of the responses received.[60] They point out that 'a substantial proportion of the queried organizations, whether out of inability or out of unwillingness, are non-compliant with the law'.[61] Sometimes it takes a 'follow-up request to receive an answer with data that was previously withheld', and 'while many replies are quite elaborate, even these replies frequently provide inadequate information to the individual for making an informed judgment about the lawfulness of the processing'.[62] Most importantly, despite the fact that a right to access has been in place for over fifteen years, some large organizations processing personal data reported that they had never received an access request. As a way forward, this study argues that 'collective use of the right of access can help shift the power imbalance between individual citizens and organizations in favour of the citizen'.[63] Along a similar line, Veale and others discuss both the desirability and challenges inherent in the development of some automated platform 'to enable data subjects to utilize their rights'.[64]

As a concrete way of addressing the above concerns, our data Trust proposal hinges upon the possibility of assigning those seldom-used rights to a data trustee, who would exercise them on behalf of the Trust's beneficiaries (and settlors). This assignability question is discussed below.

The possibility (and advantages) of holding data rights under a legal Trust

A legal framework report entitled 'Data Trusts: legal and governance considerations'[65] was recently

55 The limits inherent in the right to portability are further discussed in Helena Ursic, 'Unfolding the New-Born Right to Data Portability: Four Gateways to Data Subject Control' (2018) 15 SCRIPted 42.

56 Art 20 GDPR requires data to be available in a commonly used, structured, and machine-readable format.

57 It is interesting to note that the scope of the CCPA's portability obligations are arguably broader: if the information provided in response to an access request is delivered electronically (rather than by mail), it must be 'in a portable and, to the extent technically feasible, in a readily usable format that allows the consumer to transmit this information to another entity without hindrance'.

58 Art 16(4) of DIRECTIVE (EU) 2019/770 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 20 May 2019 'on certain aspects concerning contracts for the supply of digital content and digital services' reads: 'Except in the situations referred to in point (a), (b) or (c) of paragraph 3, the trader shall, at the request of the consumer, make available to the consumer any content other than personal data, which was provided or created by the consumer when using the digital content or digital service supplied by the trader. The consumer shall be entitled to retrieve that digital content free of charge, without hindrance from the trader, within a reasonable time and in a commonly used and machine-readable format.'

59 At the moment Recital 59 of the GDPR recommends that organisations 'provide means for requests to be made electronically, especially where personal data are processed by electronic means'. Interestingly, for our purposes, Recital 63 advises that, where possible, organisations should be able to provide remote access to a secure self-service system which would provide the individual—or in our case the Data Trustee on behalf of the data subject—with direct access to his or her information.

60 Interestingly, this study highlights that 'in most cases a request for information about specific data in an initial data request is ignored, while follow-up requests get an individualized reply more often. Sometimes a follow-up request does receive an answer with data that was previously withheld'. RLP Mahieu, H Asghari and M van Eeten, 'Collectively Exercising the Right of Access: Individual Effort, Societal Effect' (2018) 7 Internet Policy Review. doi: 10.14763/2018.3.927.

61 Ibid.

62 Ibid.

63 Roger Brownsword, 'Knowing Me, Knowing you - Profiling, Privacy and the Public Interest' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen: Cross-disciplinary perspectives* (Springer 2008).

64 M. Veale and others, 'Automating data rights' in D Eyers and others (eds), *Towards accountable systems* (Dagstuhl Reports, Schloss Dagstuhl 2018).

65 Reed, BPE Solicitors, and Pinsent Masons, 'Data Trusts: legal and governance considerations', April 2019, https://theodi.org/wp-content/uploads/2019/04/General-legal-report-on-data-trust.pdf.

commissioned by the Open Data Institute (ODI). This report claims that:

> data is not capable of constituting property in the legal trust sense, and thus cannot form the basis of a legal trust in any of the legal systems which have a concept of trust law.[66]

In what follows we will outline why this assertion betrays a misunderstanding of the nature of equitable property. To do so, we survey the current debate about the extent to which some kinds of data may plausibly be said to give rise to property rights. The latter debate is considered for interest's sake, as our data trusts proposal does not hinge on this property rights debate. We subsequently outline the advantages of relying on Trust law as a legal framework (in contrast to the contractual or corporate frameworks proposed in the above-mentioned report), as well as its challenges (rights assignability prime among them).

Data rights as the subject matter of the Trust

There are several ways of explaining what led the authors of the above report to state that 'data is not capable of constituting property in the legal trust sense'. The authors of that report could have been referring to the fact that data is an intangible asset.[67] Some jurisdictions' understanding of property (Germany most notable among them) does not allow the latter to include 'non-tangible objects'. The intangible nature of data is however unlikely to have been the driving concern behind the report's conclusion, since in the UK property rights can be established on any type of tradeable assets[68] (tangible or intangible[69]), and bank accounts are commonly held in Trusts. In the latter case, the trustee holds a personal right to payment ('right in personam') against a bank,[70] in trust for the beneficiary.

Alternatively, the authors of the report may have been referring to the fact that data differs from assets like bank balances in several respects, key among which is its non-rival nature: data can easily be duplicated, and this makes it difficult to exclude others from using that data.[71] In that respect, the challenge is no different from that pertaining to intellectual property rights—which are 'not dependent on any idea of there being "property" in the creative idea or endeavour'[72]—and are also commonly held under Trusts.

Most plausibly, the authors of the report may have sided with those who take the view that data cannot be said to give rise to property rights.[73] This view is opposed by those who argue that the GDPR rights to portability, erasure and access—when applicable—provide all that is needed to give rise to property rights for these categories of data. While the latter view is outlined in the next section, what follows aims to explain why the extent to which certain kinds of data may be said to give rise to property rights is neither here nor there as far as

66 Ibid.

67 As such it cannot give rise to a 'right in rem'.

68 Along this line, Sarah Worthington explains: 'English law, unlike its civilian counterparts, no longer holds to a sharp divide between tangible assets (such as land, bicycles and Picasso paintings) and intangible assets (such as shares, bonds and debts). Put in legal terms, English law no longer draws hard lines between 'property' and 'obligation' or 'property' and 'contract'. All these different types of rights are 'assets': The British Academy, The Royal Society and TechUK, Data ownership, rights and controls: Reaching a common understanding (Discussions at a British Academy, Royal Society and techUK seminar on 3 October 2018).

69 Today all sorts of intangible assets—or more precisely, the rights such assets give rise to—can be held under a Trust. Shares 'might be evidenced by a document, such as a company share certificate, but the value of the share is in the rights it gives you against the company': James Penner, *The Law of Trusts* (OUP 2016).

70 'The chief characteristic of a property right in relation to a thing is that it allows B to exclude others from making use of that thing. It has been argued that equitable property rights possess this characteristic. The most obvious problem with that argument is that B may have an equitable property right in relation to an intangible asset. For example, A can hold a personal right against Z, such as a bank account, on trust for B. In such a case, there is no independent physical thing against which B has a right. After all, there is no thing against which A has a right: A merely has a right to receive payment from Z. If I have title to land or a car I can exclude others from making use of it; it is not meaningful to speak of excluding another from making use of a debt owed to me.' Ben Mcfarlane and Robert Stevens, 'The Nature of Equitable Property' (2010) 4 The Journal of Equity 3.

71 For a critical discussion of the widespread assumption that the 'right to exclude' is central to the 'formal structure of property', see James Y Stern, 'What is the Right to Exclude and why does it Matter?' in MH Otsuka and JE Penner (eds), *Property Theory: Legal and Political Perspectives* (CUP 2018).

72 To quote Sarah Worthington's contribution to the report referred to in note 77: 'All intellectual property rights are created by statute, not by the courts. Notably, despite the "property" terminology, the protection delivered by these statutory means is not dependent on any idea of there being "property" in the creative idea or endeavour. Instead, the statute itself defines rights, and then defines remedies for their infringement, and it is these statutory rights that are then 'assets' that may be assigned or shared in all the ways that other assets can be dealt with at law.' The British Academy, The Royal Society and TechUK.

73 This interpretation is supported by the fact that the report refers to Oxford v Moss [1978] 68 Cr App 183: in the latter case the Court of Appeal concluded that confidential information (in this instance an examination paper) did not fall within the definition of 'intangible property'. Along a similar line, the report could also have referred to the more recent: *Your Response Limited v Data team Business Media Limited* [2014] EWCA Civ 281. In that case, Lord Justice Floyd stated that '[a]lthough information [in this instance, a subscribers database] may give rise to intellectual property rights, such as database right and copyright, the law has been reluctant to treat information itself as property. When information is created and recorded there are sharp distinctions between the information itself, the physical medium on which the information is recorded and the rights to which the information gives rise. Whilst the physical medium and the rights are treated as property, the information itself has never been'. Notice the distinction between the 'information' and the rights the latter may give rise to. This distinction is crucial since the subject matter of a Trust is best defined in terms of rights, not property, as explained below (in this section).

our proposal is concerned. What matters, instead, is the extent to which regulatory instruments such as the GDPR confer rights, and for what kind of data. The latter question—what kind of data gives rise to what kind of rights—will determine the possible scope of data Trusts in various jurisdictions.

Quoting Lord Shaw, according to whom '[t]he scope of the trusts recognized in equity is unlimited. There can be a trust of a chattel or a chose in action, or of a right or obligation under an ordinary legal contract, just as much as a trust of land', McFarlane and Mitchell[74] explain in their textbook that:

> If the term 'property' is not limited to physical things, and instead extends to any valuable, assignable right, the possibility of the declaration of trust of a non-assignable right [as per Don King Productions Inc v. Warren[75]] shows that the subject matter of a trust need not fall within even that extended definition of property. *For this reason, it has been suggested that it is more accurate to think of the subject matter of the trust as a right, rather than as a property*. After all, even in a case where the property held on trust is a physical thing, such as land, it is the trustees' right to the land that is held on trust, and not the land itself.[76]

Similarly, Robert Chambers seeks to dispel the 'false assumptions' made by many who 'fear the trust', and states that:

> Every trust is a relationship between at least two persons (a trustee and a beneficiary) in which the trustee holds some right in trust for the beneficiary. Almost any right can be held in trust [.][77]

If the subject matter of a Trust is best thought of as a right,[78] rather than property, the important question becomes the extent to which the rights conferred by the GDPR (and other regulatory instruments in other jurisdictions) can be held in Trust. There are no obvious public policy reasons why data rights (unlike the right to vote, for instance) should not be held in Trust, and data rights are quite different from a right—or license—to practice dentistry, say, which was deemed unsuitable given its being intrinsically connected to the right holder.[79] The latter case brings to the fore one important point: at the moment the rights to portability, erasure, and access as conferred by the GDPR are not mandatable to a third party. To make such rights mandatable (in our proposal, to a data trustee) would require regulatory intervention. This is addressed below.

The extent to which some kinds of data may be said to give rise to property rights

Today many concur in acknowledging the fact that data continues to present unique challenges when it comes to building a battery of rights and responsibilities that adequately takes into account both its growing importance as an economic asset and its human rights implications, given the social vulnerability it entails. While the GDPR set of rights and responsibilities pertaining to personal data are rooted in human rights concerns, the rights to portability, erasure and access mentioned above can nevertheless be said to provide all that is needed to give rise to what may plausibly be characterized as property rights[80] (notwithstanding obstacles specific to Italy and Germany[81]). The latter's strength varies, and decreases (all the way to nil) as we move away from data that is 'directly provided' by the data subject (1), to data such as cookies—for which there is no right to portability—(2), all the way to data that is the result of sophisticated processing, such as the data leading to credit rating scores (3).

In his 'relational taxonomy' of personal data,[82] Malgieri helpfully distinguishes between the above three levels to argue that the 'intermediate category' (2) is the most complex, calling as it does for '"shared property" based on shared exclusive rights': 'Consumers will have full exclusionary rights against all commercial actors

74 McFarlane and Mitchell also refer to Re Lehman Brothers (International) Europe, where Briggs J notes that: '[N]o-one doubts the beneficial interest of clients in a solicitor's client account. Yet the subject matter of that fund consists entirely of the solicitor's purely personal rights as a customer of the client account bank or banks'.

75 *Don King Productions Inc v Warren* (No1) [2000] Ch 291; [1999] 3 WLR 276, CA (Civ Div).

76 Ben McFarlane and Charles Mitchell, *Hayton and Mitchell: Text, Cases and Materials on the Law of Trusts and Equitable Remedies* (Sweet & Maxwell 2015), our emphasis.

77 Robert Chambers, 'Distrust: Our Fear of Trusts in the Commercial World' (2010) 63 Current Legal Problems 631.

78 As such, the beneficiaries are said to have a 'right against a right'. Mcfarlane and Stevens notably highlight the advantages of this analysis when it comes to the cross-jurisdictional applicability of the Trust law framework: 'a key practical consequence of the [right against a right] analysis is that it provides a model by which institutions such as the trust or equitable assignment can be accommodated within legal systems that have not experienced the productive paradox of two rival court systems [Equity and common law].' Mcfarlane and Stevens 9.

79 *Caratun v Caratun* (1992) 96 DLR (4th) 404, 10 OR (3d) 385 (CA).

80 Purtova, Andreas Boerding and others, 'Data Ownership—A Property Rights Approach from a European Perspective' (2018) 11 Journal of Civil Law Studies 323. In an American context, see also the earlier work of Schwartz who defends the need for a 'qualified 'propertization' of personal data: Paul M Schwartz, 'Property, Privacy, and Personal Data' (2003) 117 Harvard Law Review 2056.

81 In Germany, s 90 of the BGB specifies that only physical objects can constitute 'things' in the legal sense (ie be the object of property rights). Similarly, art 810, together with art 814, of the Italian Civil Code specifies that data cannot be a 'good' given its intangibility. The latter class of assets calls for a special law if it is to give rise to property rights.

82 Gianclaudio Malgieri, 'Property and (Intellectual) Ownership of Consumers' Information: A New Taxonomy for Personal Data' (2016) 4 Privacy in Germany PinG 133.

interested in their data (including the company which has a shared ownership on such data); the data controllers/businesses will be able to exercise their exclusionary rights against all competing companies but not against the data subject (whose interests prevail).' In the weakest category (3), in contrast, the extensive processing needed to give rise to data such as credit rating scores entails that the 'quasi-property of companies prevail on control rights of individuals'.[83] The rights individuals do have in relation to such 'weakly relating data' (3) do not have any of the characteristics of property rights: they are best characterized as consumer protection rights, such as Article 14 GDPR's right to information about the controller, processing, purpose, etc.

Challenges and opportunities inherent in holding data rights under a Trust (rather than a contractual agreement or corporate structure)

On the challenges front, the varying extent to which different kinds of data can be deemed to give rise to different levels of property rights—from full portability, access and erasure rights all the way to nil (ie 'mere' information rights)—is a problem only if one assumes that the subject matter of a Trust must be able to be defined as 'property', an assumption which the above section has sought to dispel.

Other difficulties include the fact that it can be difficult to ascertain what data relates to which identifiable person: sometimes data relates to more than one person at a time (shared provenance issues are discussed in the section discussing implementation challenges), sometimes the extent to which data is 'identifiable' varies over time, depending on the degree of data aggregation.[84]

The weight of the challenges stemming from these identifiability issues (and the non-rival nature of data) will in part depend on specific implementation choices, which are discussed below. For now it is worth emphasizing at this stage that a data Trust does not necessarily have to 'pull' the data held by various data collectors. One can imagine an ecosystem of Trusts where one Trust A specializes in direct data management (possibly choosing to locate its servers in a specific jurisdiction), while Trust B devolves responsibility for data management to Trust A. Trust B would then focus on the policy, rather than the practicalities, underlying data sharing. This would allow trustees to focus on the use to which the data is put rather than the detailed mechanisms of access and storage which would be standardized. Alternatively, Trust C may work on the basis of a wholly decentralized model, whereby the beneficiaries' data stays wherever it is. Depending on the particular model adopted, data trustees may or may not be deemed data controllers under the GDPR.[85] Any Trust may choose to share data with other Trusts that conform to their constitutional terms. While some Trusts may be set up to manage and protect as much of the data pertaining to their beneficiaries as possible, other Trusts may specialize in only a particular kind of personal data, such as health data. Such specialized Trusts are likely to want to negotiate with the more generalist Trusts so as to be able to reap the benefits that come with large-scale datasets.

Most significant among the challenges mentioned so far is the possibility of mandating the rights to portability, access and erasure mentioned earlier. In Europe, Article 80(1) of the GDPR reads:

> The data subject shall have the right to mandate a not-for-profit body, organisation or association which has been properly constituted in accordance with the law of a Member State, has statutory objectives which are in the public interest, and is active in the field of the protection of data subjects' rights and freedoms with regard to the protection of their personal data to lodge the complaint on his or her behalf, to exercise the rights referred to in Articles 77, 78 and 79 on his or her behalf, and to exercise the right to receive compensation referred to in Article 82 on his or her behalf where provided for by Member State law.

The above wording has given rise to a certain amount of controversy, since it is debatable whether the 'where provided for by Member State law' applies to the whole sentence or to the right to receive compensation only.[86] For our purposes, what matters is that Article 80(1) only mentions the rights in Articles 77, 78, and 79 as

83 Ibid.

84 'The same piece of data, depending on a particular context, can be personal and non-personal, more or less likely to relate to an identifiable natural person, and with a stronger or weaker link to that person. [...] The difficulty lies, first, in determining at which point the level of relation to an individual is sufficient to establish property rights, and second, in tracing the presence of such a relation.' N Purtova, 'Do Property Rights in Personal Data make Sense after the Big Data Turn: Individual Control and Transparency' (2017) 10 Journal of Law and Economic Regulation 64.

85 It is beyond the scope of this article to discuss the scope and nature of the legal obligations stemming from a data trustee's potential 'controller' status.

86 For a summary of this controversy, see Alexia Pato, 'The National Adaptation of Article 80 GDPR: Towards the Effective Private Enforcement of Collective Data Protection Rights' in K Cullagh, Olivia Tambou and S Bourton (eds), *National Adaptations of the GDPR* (Collection Open Access Book, Blogdroiteuropeen 2019) n 9. Pato notably refers to the wording of the Italian version, the historical development of art 80 and the contrasted wording of art 80(2) to conclude that only the possibility to extend the representative action to the right to compensation is left to the discretion of the member States.

'mandatable'. This raises the following question: would national legislation aimed at making the rights to portability, erasure and access mandatable (to a data trustee) be in breach of EU law? A positive answer is unlikely, since such intervention would not necessarily extend (nor diminish) the scope of controllers' obligations. Given the current, well-documented difficulties in exercising the rights to access, portability and erasure, there are pressing, positive reasons for regulatory intervention on this front. It is also worth emphasizing that this mandatability issue would be just as much a problem for the contractual or corporate structures envisaged in the 'Legal Framework report' mentioned earlier,[87] insofar as they also rely on an appointed third party to exercise the rights to portability, access, and erasure.

Now, while the above challenges are by no means negligible, they do not constitute reasons to doubt that data rights can held under a legal Trust. To conclude otherwise and 'make do' with a contractual or corporate framework is not only to impose an unjustifiably narrow understanding of equitable Trusts and their purported subject-matter. It is also to deprive society of a particularly valuable governance tool given the unprecedented challenges (and vulnerabilities) at stake. Not only is a Court's equitable jurisdiction to supervise and intervene if necessary not easily replicable within a contractual or corporate framework, the importance of the fact that '[e]quity employs ex post moral standards, emphasizes good faith and notice, couches its reasoning in terms of morals, and is sometimes vague rather than bright line'[88] cannot be overestimated.

Case studies

The following 'case studies' are meant to illustrate the complementarity of our data Trusts proposal with the legal provisions pertaining to different kinds of personal data. This delineation into kinds is for illustrative purposes only—most data Trusts are likely to encompass all or at least several of these 'categories' of data (which may overlap).

Medical data

Patient consent to participate in specific medical studies is currently undertaken on an individual basis, often confronting patients with a significant decision that must be taken at a moment of greater vulnerability, faced with the imminent diagnosis of a potentially serious disease. This individual consent determines who may have access to the data shared by the patient, and for what purpose. Because the terms by reference to which consent is given vary across different studies, it is often difficult to pool data obtained in the context of different studies, even if the patient's intent was to share their data more generally with organizations that bring benefit to the wider public. The GDPR accounts for this challenge, notably by relaxing the purpose specificity requirement. To take into account the fact that data mining techniques often search for correlations within disparate datasets, recital 33 suggests that the purpose of data collection can be defined very widely, merely referring to a certain area of research.

The emergence of an ecosystem of data Trusts could go further still in removing current obstacles to research while at the same time improving data subjects' ability to make choices that reflect their aspirations. The need to choose among different Trusts would indeed encourage patients to think about their sharing preferences *before* possibly being placed in a vulnerable position. The necessity to consider the specific requirements of clinical studies—and different ways of accommodating those requirements within different approaches to data governance—would also be more transparent and thus amenable to much-needed societal debate.

As an example of such requirements, consider the need for a clinical study's samples to be randomized: maintaining the validity of a study's conclusions indeed often requires continual access to consented data, at least over a given time period. Again, the GDPR addresses this requirement by allowing researchers to further process personal data for research purposes in spite of a data subject's request for erasure, insofar as this request is 'likely to render impossible or seriously impair the achievement of the [research] objectives' (Article 17(3)(d)). Similarly, a researcher may override a data subject's objection to processing if 'the processing is necessary for the performance of a task carried out for reasons of public interest' (Article 21(6)). As for what counts as 'public interest', it is left open to further specification: Recital 45 merely specifies that it 'should have a basis in Union or Member State law'.

Given its complexity, the legal framework hinted at in the above paragraphs is unlikely to be grasped by even the best-informed patients. In contrast, data trustees may meaningfully balance their Trust's commitment to providing data for societal benefits—if any—with the terms governing data sharing. They may achieve such balance by for instance making sure that access to the data is restricted to health professionals, with particular

87 Reed, BPE solicitors and Pinsent Masons (see n 1).

88 H Smith, 'Property, Equity and the Rule of law' in L Austin and D Klimchuk (eds), *Private Law and the Rule of Law* (OUP 2014).

safeguards in place. Once a significant number of people join particular Trusts, the relevant data trustees may well be in a position to negotiate safeguards that go beyond (or substantiate) those currently specified by the GDPR—Article 89 (1)—for research purposes. At the moment, the GDPR requires that technical and organizational measures be put in place to ensure that only the personal data necessary for the research purposes is processed (in accordance with the data minimization principle (Article 5(c)). Recital 33 also rather vaguely states that the processing must be 'in keeping with recognized ethical standards for scientific research'. Given the contested nature of such ethical standards, data trustees may be brought to play a significant role in the debate that currently surrounds 'ethical standards' when it comes to data research. They are also more likely to be in a position to monitor compliance with such safeguards, and possibly steer what is considered 'best practice' within particular research studies.

Social media data

As we progressively appreciate the importance of the social and environmental determinants of health, any endeavour to neatly delineate what counts as 'health data' versus 'social media data' is increasingly futile: social media interactions can be indicative of both our mental and physical health. Simultaneously, our interactions on social media change our perception of the world around us in ways that we do not directly control. The adverts we are shown and the structure of our online environment (including information feeds) are determined by algorithms that seek to maximize user engagement. They can be validated through large-scale A/B testing, monitoring in real time the effect of particular adverts (or news content) on the level of engagement shown by users with particular profiles. As users' attention is more likely to be grabbed by content that reinforces their existing preferences, beliefs or fears, the drive to maximize user engagement not only leads to 'filter bubbles'. It also increases the extent to which users are likely to indiscriminately accept fake news.[89]

For an individual to make a particular choice about how data from their social media feed is shared when they interact with these sites may require examination of extensive terms and conditions each time they join a site. Additionally, if terms and conditions change they may require re-examination. Since social media platforms often provide valuable tools for communities, users may feel obliged to accept any changes to data processing terms and conditions for fear of losing these benefits.

These power asymmetries would be addressed by data Trusts. As more people join data Trusts, terms and conditions negotiations would be handled by each data Trust. Rather than stipulating whether a user agrees to particular terms and conditions, users would simply state which data Trust they belong to. Their data could then be dealt with accordingly. This negotiating power is not a structural condition of the proposal: it is merely a side effect of the power that would accrue to the Trust through the pooling of data rights.

Genetic data

Genetic data presents particular challenges because our genome encodes not only information about ourselves but our relatives too: sensitive information can leak through other individuals sharing their genomic data. While capturing historic serial killers[90] may be unambiguously seen as a good thing for society, other details can leak through genetic data, such as misallocated parenthood. These issues are sensitive and personal. Even those who argue that a child always deserves to know the truth about their parentage would readily acknowledge that such information should be revealed sensitively to the individual concerned, not accidentally via their siblings or distant cousins.

Like all personal data whose provenance is shared,[91] genetic data does not lend itself to 'standard' access, portability, and erasure rights. While the inclusion of 'genetic data' within the GDPR's 'special category' (Article 9) does acknowledge the sensitivity of the rights at stake, the delineation of adequate safeguards—given the research exemption in Article 9(2)(j)—is largely left to national legislators, prompting some to worry about their effectiveness.[92] Our data Trust proposal is not meant to 'solve' all the issues surrounding data whose provenance is shared. What it can do is provide a sorely needed 'bottom-up' forum for societal debate, and possibly point towards new ways of approaching those problems. Genetic data could, for instance, lend itself to

89 M Mitchell Waldrop, 'News Feature: The Genuine Problem of Fake News' (2017) 114 Proceedings of the National Academy of Sciences 12631.

90 Over the last ten years, familial DNA searching (whereby DNA voluntarily uploaded to a genealogy or family database is relied on to find close matches for unidentified DNA evidence, in which case users of the genealogy database may unwittingly become 'genetic informants') is increasingly being used to solve 'cold' cases on both sides of the Atlantic.

91 This shared provenance issue applies in a range of contexts, from ambient monitoring or surveillance (image and/ or sound) to social media feeds.

92 Kärt Pormeister, 'Genetic Data and the Research Exemption: is the GDPR Going too Far?' (2017) 7 International Data Privacy Law 137.

the formation of a specific, emergent form of data Trust, where there is no longer direct overlap between beneficiaries and settlors. As a settlor I may indeed 'entrust' my genomic data for the benefit of my children (or other relatives). Depending on the terms of this familial data Trust, a scientific organization may or may not have access to this data under certain conditions. Such a 'familial' model may also be applicable to other forms of data where personal provenance is shared (see note 45), even if the origin is not familial relationships.

Financial data

Accumulated financial data (in combination with other datasets) is used to make decisions that range from credit-worthiness to identity verification. In the past, the credit-worthiness of an individual or company would have been confirmed by a letter from their bank manager or an audit of corporate accounts. Today, banks centralize financial transaction information with credit bureaux (or in the UK, credit reference agencies), who then validate an individual's credit worthiness. Such validation is normally a prerequisite for obtaining a loan or credit card. Without this information, lenders would have to rely on self-declaration of financial status, which would leave them exposed to dishonesty.

While the use of credit agencies rectifies the information asymmetry between individuals and lenders, it creates a power asymmetry between credit reference agencies and data subjects. An individual data subject is required to comply with the stipulations of credit bureaux to receive a loan, but the subject has little to no representation in steering those stipulations. Regulatory interventions can only go so far in addressing the challenges raised by this power asymmetry: in the UK the credit bureaux are commercial entities regulated by the Financial Conduct Authority. They have responded to the GDPR with the Credit Reference Agency Information Notice, which outlines how each agency uses and shares personal data, and for what purposes. Unsurprisingly, erasure rights are severely restricted, as is the ability to object to further processing: in both cases, there is a strong likelihood of being overridden by the agency's 'overriding legitimate interest to continue processing' (data portability rights do not apply, given the reliance on 'legitimate interests' as the ground of processing). Given these limitations, do data Trusts have any role to play here?

We believe they do. Data Trusts could leverage the right to information (Article 14 GDPR) of their members to ensure greater transparency in the operation of the credit reference agencies. This would enable the data trustees to ensure that care of the data subjects' personal data is not compromised by the commercial interests of the Credit Reference Agencies. Breaches of Data Protection do occur: Equifax[93] Ltd was recently given the maximum possible fine of £500,000 under the 1998 Data Protection Act. But this is after the fact law enforcement. From a more constructive perspective, Data Trusts might provide a mechanism to ensure that the data subjects' voices are better heard in the drafting of data sharing terms, and in ensuring best standards are adhered to.

Loyalty card data

Loyalty programs encourage consumers to continue to shop at the same outlets by rewarding repeat visits or purchases. Loyalty programs have evolved to also better characterize each consumer, thus allowing targeted advertising. Like social media data, there is great potential for loyalty card data, particularly from supermarkets, to be used in the context of personalized health analysis. Loyalty card data could for instance provide information about what individuals or families have been consuming in their diet: traditionally, dietary intake information is gathered by self-reporting. Yet, evidence suggests that this self-reporting approach is inaccurate, with biases towards perceived norms.[94]

Since such potential medical uses are not normally anticipated in the terms and conditions of the Loyalty card scheme for which individual consent is required, 'digital receipts' have yet to be used extensively in medical studies. In contrast, a data Trust could stipulate in advance that loyalty card data should be made available for medical research, under certain conditions. Thanks to such a data-portability stipulation, medical studies would be able to obtain easier access to this potentially fruitful data from a number of different loyalty card schemes.

Implementation challenges

First, it is worth emphasizing at the outset that our data Trust proposal would be able to reach across different

93 On 7 September 2017, Equifax, a US-based credit monitoring company, announced that over 140 million consumers' personal information had been stolen from its network. The subsequent ICO investigation found that Equifax's UK arm had not taken the necessary steps to ensure that Equifax Inc—the American parent company which was processing consumers' data on its behalf—was adequately protecting consumers' personal information.

94 Dale A Schoeller, 'How Accurate is Self-Reported Dietary Energy Intake?' (1990) 48 Nutrition Reviews 373; BM Appelhans and others, 'To what Extent do Food Purchases Reflect Shoppers' Diet Quality and Nutrient Intake?' (2017) 14 The International Journal of Behavioral Nutrition and Physical Activity 46.

jurisdictions, despite the Trust's historical origin: Trust structures find their roots in the development of the 'court of equity' in 14th-century England. The 'court of equity' was born out of the need to provide remedies to claimants—such as returning crusaders who had transferred the title to their land while on crusade—when none were available under the common law: the spirit underlying our 'data Trust' proposal is not dissimilar, in that the 'remedies' currently provided to data subjects can be seen as deficient. Today, Trust structures can and do operate in non-common-law jurisdictions,[95] and there is a growing interest in their cross-jurisdictional aspects: many social and legal functions of a Trusteeship are served by analogous 'offices' in civil law,[96] if by other names.[97]

Among the challenges that will have to be critically considered for our proposal to become a 'live' possibility (and lend momentum to budding initiatives[98]), two deserve a special mention: uptake and exit procedures.

To start with uptake: the novel and multifaceted nature of the risks pertaining to personal data is difficult enough to grasp for the 'actively interested' individual who is computer literate. Many people are not even aware of the fact that most of the personal data held by corporations is 'passively' obtained through ambient tracking devices. As a result, the average level of interest in registering with a data Trust may be low. If data Trusts are not to end up as a means of increasing the bargaining power of only the least vulnerable part of the population (ie those that are already data-aware), a variety of measures ought to be considered. The latter could range from 'simply' compelling large data controllers to flag up the existence of a variety of data Trusts and their underlying benefits, to possibly implementing some 'default' data Trusts—focusing for instance on local data sharing needs: in the absence of choice, data subjects would be assigned to such 'default', publicly funded Trusts, with frequent, proactive reminders about the possibility of joining alternative Trusts. From a justificatory perspective, such a default policy could find its roots in reasons that are very similar to those that have led to the default provision of a pension fund (ie a poor understanding of the long-term risks impairs an ability to make informed decisions). Yet, the dangers inherent in such a paternalist approach warrant great caution: aside from the need to make sure that opt-out procedures remain extremely accessible throughout (and robustly implemented), one would also need some ongoing review process. The latter would not only nudge those who have been assigned to a default Trust to consider switching, but also review the extent to which the terms of the default Trust do optimally serve the needs and aspirations of its beneficiaries.

A related challenge stems from security concerns. One may worry that data Trusts may inadvertently increase the extent to which the data that is collected ambiently can be traced back to particular individuals. Take the rotating MAC address currently used by iPhones to minimize the extent to which our passive data trail can be traced back to particular users as an example: does the emergence of data Trusts structures mean that such protective measures would have to be relinquished?[99] Not necessarily: a variety of differential privacy techniques may be relied on to address such risks.

Another challenge relates to exit procedures: many of the 'smart' devices and appliances collecting user data are used in a way that makes it very difficult, if not impossible, to find any data that is related to one user only. In that context, how does one determine what is owed to a person leaving a particular data Trust? This question is one that current data controllers are already familiar with and each Trust may specify different ways of disentangling data for the purpose of exit procedures. A related concern bears upon accountability procedures for data trustees. Their being held to the high standards entailed by fiduciary duties may not make that much of a difference if the data Trustee is unable to compensate for the harm created by lax data management. Should data Trustees hold liability insurance, and if so who pays for it? Can the public nature of the services provided by data trustees (and the vulnerabilities they address) be deemed similar enough to those provided by medical doctors to justify their being publicly funded? Should data Trusts be overseen by a regulatory body that would set training requirements and some code of

95 See also note 71 on the cross-jurisdictional applicability advantages of the 'right against a right' analysis referred to in section ' Data rights as the subject matter of the Trust'.

96 There is growing interest in those aspects of Trust Law that can plausibly be imported into domestic civil law, and in the harmonization potential that would result from such efforts: Reinout Wibier, 'Can a Modern Legal System Do without the Trust?' in Lionel Smith (ed), *The Worlds of the Trust* (CUP 2013) and Ruiqiao Zhang, 'A Comparative Study of the Introduction of Trusts into Civil Law and its Ownership of Trust Property' (2015) 21 Trusts & Trustees 20; Raúl Lafuente Sánchez, 'Recognition of Foreign Trusts and Challenges Facing the Spanish Courts' (2017) 23 Trusts & Trustees 12.

97 Civil legal systems have traditionally denied the domestic applicability of Trust Law, given its divergent underpinnings—and its ability to be deployed to skirt the law.

98 Existing personal data repository initiatives such as 'Citizen me', 'midata', etc would benefit from the development of an ecosystem of Trusts.

99 Mavroudis and Veale (n 2).

conduct for data trustees, together with an expert advisory body tasked with scoping out the long-term problems and/or side-effects associated with particular types of data governance?

Conclusion

The legal institution of Trusts was born—almost 700 years ago—out the lacunae of the Common Law: the latter was for instance unable to provide remedies to those who had trusted others with the title to their land while on crusade. The problems which our data Trust proposal seeks to address do not have much in common with those of 14th-century crusaders. Indeed, it seeks to *reverse*—rather than perpetuate—a data governance framework that is strikingly similar to a feudal system, whereby data subjects' leaked data is exploited by increasingly large data controllers in a seemingly inexorable way. Laudable as they are, current regulatory endeavours to curb contractual freedom cannot by themselves reverse those power imbalances. Nor can they suffice to address the slow insidious compromising of our ability to maintain a social self that is at least partly controlled by us. Remedies for the latter ills are unlikely to be found exclusively in further, 'one-size-fits-all,' top-down regulation.

Our data Trust proposal aims to empower us, data subjects to 'take the reins' of our data in a way that acknowledges both our vulnerability and our limited ability to engage with the day-to-day choices underlying data governance. The availability of a variety of data Trusts—each reflecting a particular set of aspirations (and attitude to risk)—not only promises a degree of adaptability that top-down regulation is unlikely to match. It is also conducive to a much greater level of societal awareness and debate. As a vehicle facilitating the constructive articulation of data governance aspirations, an ecosystem of data Trusts addresses needs that are complementary to those within the reach of regulatory interventions such as the GDPR (including collective enforcement aspects). Importantly, by potentially facilitating access to 'pre-authorized', aggregated data (consent would be negotiated on a collective basis, according to the terms of each Trust), our data Trust proposal may remove key obstacles to the realization of the potential underlying large datasets.

To be effective, the Data Trusts we propose need to be representative of the data subjects concerns. A successful data Trust will be one whose constitutional terms better encapsulates the aspirations of a large part of the population. That Trust would, in turn, yield more influence over data controllers. This ascendancy, combined with the fiduciary responsibility of the data Trustees, is key to rebalancing power imbalances within our current system of data governance. Seeding an ecosystem of data Trusts (ideally through a combination of public and private initiatives), together with the creation of a body of competent data Trustees, is a key component to bringing about such rebalancing.

doi:10.1093/idpl/ipz014