Noname manuscript No.
(will be inserted by the editor)

Liberalism and the digital body

Rethinking autonomous agency in data protection

Julia Netter

Received: date / Accepted: date

Abstract The shift of everyday activities to online platforms and services, and the plethora of digital data that a modern person generates both implicitly or explicitly as they go about their life poses new challenges to individuals' capacity to reflect on, and make meaningful decisions with regard to, the information they provide about themselves. In this paper, I argue that common liberal paradigms like consent and targeted regulation when applied to digital technology fail to provide meaningful orientation on how to approach the challenges of data protection. I claim that for liberal theory to determine what it means for a person to be an autonomous agent in the data-saturated online world, we must expand our conceptual devices, and, in particular, build upon a more nuanced view on what constitutes the individual person in the digital sphere. I develop the notion of a digital body, the collection of data that an individual creates deliberately and implicitly, and which – akin to a physical body – is a medium for others to act on the individual. The digital body allows us to reason about the prerequisites for and dangers to autonomous agency in the digital sphere, and provides us with a starting point for deliberating adequate means for protecting it.

Keywords data protection · autonomy · liberalism

1 Introduction

Liberal states and societies approach the digital sphere with the paradigms of the analogue world. The most common response to challenges of privacy and data protection online relies on the tools of consent as well as targeted

Julia Netter
Brown University
E-mail: julia_netter@brown.edu

regulation, which are both rooted in a broadly liberal commitment to individual autonomous agency, attempting to carefully balance the freedom of online entrepreneurs and users, of content creators and consumers.

The whole ecosystem of online services funded by personalised advertising currently relies on users obliviousness about the scope and interconnectedness of individuals' digital selves. For these businesses to flourish, it is indispensable that users keep providing their personal data. But instead of choosing to provide information about themselves and others that they are comfortable with, users are simply unaware of the extent and content of that data and how it might ultimately be used to influence or shape their actions either through manipulative practices (*e.g.,* in advertising, political campaigns) or coercion (*e.g.,* by law enforcement).

My aim is to convince you that the paradigms of consent and regulation as they are currently employed to data regulation are inadequate in light of the transformative changes brought about by the data economy. This is because the shift to living a large portion of our economic and social lives online not only changed the platforms and formats of our interactions with others, but it fundamentally changed the shape and boundaries of the individual person: I argue that it furnishes us with additional *"digital bodies",* consisting of the personal data that is tied to us as individuals, both in terms of the information we actively provide and that which is generated from our actions online. These trails of data are 1) ubiquitous, 2) invisible and intangible, for the vast majority of people, 3) interconnected, *i.e.,* one persons data often contains or allows inferences to information about others.

These characteristics fail to be reflected in how liberal states respond to challenges and abuses in the new data-driven society. In section 2, I demonstrate that relying on individuals explicit authorization for the collection and analysis of their data – drawing on a literal interpretation of autonomous self-determination online – is meaningless. The supposed consent of users is undermined by the near-constant need for individuals to make such decisions as well as the intangibility, and complexity of the choices provided. Furthermore, I argue in section 3 that it is unclear if liberal societies currently posses the justificatory resources for presenting a compelling case for targeted regulation of data-collection practices since these practices cannot easily be placed on the continuum of already existing regulatory practices in other areas, mainly because individuals are unable to clearly see and meaningfully reflect on the benefits, risks, and harms of data protection regulation.

In short, I make the case that individuals' capacity to make meaningful autonomous choices about data protection and privacy easily fall through the cracks of liberal devices like consent and targeted regulation. In section 4, I explore how approaching individuals' relation to the data they generate in the digital sphere through the concept of the *digital body* helps us pinpoint the specific reasons for the failure of the paradigms of consent and regulation in their current form and even provide the justificatory basis for a new principled approach to regulation.

2 Consent is broken

Every one of us grants consent to data use on the Internet on a regular basis. We do it when we accept the terms of service and privacy policies of a web service, and we do it at small scale when we dismiss the cookie banners that websites deploy to comply with requirements of the European Unions General Data Protection Regulation (GDPR). Yet, in practice, very few of us actually know precisely what we consent to, even though legally we are providing active, informed consent.

Consent is an elemental concept in liberal political philosophy.[1] One of the foundational tenets of modern liberal societies is to assume in a normative sense that persons are autonomous agents capable of self-directed action (see, *e.g.,* Raz 1986, 371; Waldron 1993, 155-156). Generally speaking, respect for autonomous agency assigns to individuals a prerogative to determine the shape of their own lives – their commitments and actions. On a very abstract level, this means that individuals get to decide what others are authorized to do to them or to extensions of themselves, such as their property.

In the offline world (*e.g.,* in medical settings), active, express, informed consent is the gold standard for giving others permission to act on us.[2] In online services, requiring users to consent to terms and conditions is a widely used tool for the authorization of the use of data as well.

However, active express consent on the Internet is broken. Users are asked to read through long and complicated pages of legalese, which is both tedious and often unhelpful in making people understand what precisely they are asked to consent to. In addition, required consent to terms and conditions is ubiquitous in the digital realm and the stakes are usually low (in the eyes of the users), so people just click to agree to the terms. Therefore, users consent in the vast majority of cases is not meaningful, as they do not even consider, let alone understand, what they are consenting to. From a normative perspective, the value of such expressions of "consent" is dubious at best.

Prior technical approaches to consent, developed in computer science research, seek to provide users with better explanations of what they are consenting to, or seek to make consent requests fine-grained and rooted in the context of an explicit user action (*e.g.,* an app activating the camera on their phone). While valuable, these approaches are insufficient on their own. Improved explanations — through visualization (Kitkowska et al. 2020), clear language, or standardized user interfaces — break down dense legalese, but simplify the proposed data use sufficiently to nevertheless impede meaningful consent. For example, a broad statement of we may use your data to show you ads lacks information about the specific data used and the type of processing, while highly detailed and technically accurate explanations will overwhelm most users. In practice, users willingness to consent depends heavily on how

[1] This is reflected most prominently in social contract theory. See, for example. Locke (1988), Rawls (1972).

[2] See, for example, the Nuremberg Code which was instrumental in enshrining informed consent as a basic principle of medical research (Shuster 1997).

the request is presented (Wei et al. 2020). Moreover, a permanent consent lacks meaningfulness if the data use changes over time, while repeated consent requests desensitize users. Another approach, contextual consent requests, as in recent smartphone operating systems (Tsai et al. 2017), makes users aware of the likely purpose of data use: if an app asks to use the camera in the context of sharing a photo, the user intuits the purpose from context. However, most service providers use backend processing that is invisible to the end-user (*e.g.,* training AI models for targeted advertising), and contextual consent mechanisms cannot be used for these non-interactive data uses (nor would end-users likely understand them).

In short, users' consent to the processing of their data is not meaningful, because, in practice, people do not read what they are asked to consent to and, even if they did, they are unlikely to realize precisely what kind of data they are asked to give up and what implications giving it up might have for them. To argue that data processing only requires more transparency and more fine-grained opportunities for user control is to hide behind an essentially vacuous commitment to individual user autonomy. What is more, holding up consent and user choice as the key paradigm for shaping and regulating the digital sphere also serves to diminish wider debates about the harms and benefits that come with the ubiquitous collection and processing of large amounts of user data online. After all, from a liberal perspective, enabling individuals' capacity to choose when and how much data they want to give up, and in which context, looks like the gold standard. Relying on individual consent is attractive if we assume that individuals are best placed to look out for their own interests. Yet, as it turns out, when it comes to data processing, individuals are mostly unable to look out for their interests, because the scale, complexity, and ubiquity of the choices they face makes it nearly impossible even for highly informed users to discern at every point what precisely they are asked consent to, what the potential implications of their consent are, and if they care about them. So, if consent online is broken, how can we ensure that individual users are protected from the potential harms that come with the collection and processing of the vast amounts of data that they themselves hand over to online service providers on a daily basis?

3 What about regulation?

A logical next step would be to move towards imposing actual limitations on how online service providers can collect, store, process, and transfer their users data, instead of merely requiring them to ask for their users' consent. This kind of data protection regulation is essentially paternalistic. Interfering with, broadly speaking, the relationship between service provider and users is, in this case, not justified by potential harms to others but to the users themselves.[3]

[3] There may be collective harms, and to the extent that there are, this will make regulation easier to justify. But for the purpose of this paper, I restrict the scope of the analysis to harms that merely affect the individual who provided the data.

If, for example, data protection regulations were to prevent an online service provider from transferring their users' data to a large number of other service providers, such as for the purpose of improving targeted advertising, in order to protect users from data leaks and subsequent dangers of fraud and identity theft, we essentially allow those who design the regulation to make a decision for individual users that we think (for good reason, see above) these users are not very well placed to make for themselves.

Many instances of regulation are broadly paternalistic by that standard: they ultimately exert influence over the range of options available to the consumer – eliminating or limiting the availability of harmful products and services – for his or her own benefit. Take, for example, food safety regulations, which, starting in the 20th century, worked to eliminate known toxic substances (*e.g.,* lead and arsenic) from processed food products (Blum 2018). Regulation of this kind takes advantage of the supposedly superior capacity of the state – compared to the individual consumer – to collect and process information on the prevalence and detrimental effects of harmful substances and practices, and uses its power to enforce standards which benefit, and are generally aligned with, the obvious interests of consumers where individuals are in no position to make those choices. Regulation banning the use of formaldehyde as a preservative in milk provides consumers with a benefit (reduction in their exposure to toxic substances) which is clearly in their interest, without altering the fundamental character of the product they want to consume (milk that is less toxic is still milk). Data protection regulations which require online service providers to use an adequate standard of encryption for their databases in order to protect their consumers' data from being compromised in a hacking attack by malevolent actors are of that kind.

Some instances of data protection regulation are paternalistic in a more narrow sense – and more in line with the colloquial use of the term – in that they also narrow choices which are actually attractive to at least some consumers and would otherwise be available to them. If a state were, for example, to ban or restrict the consumption of red meat (by means of taxation, licensing, or selective prohibition) because it has been shown to increase the risk of heart disease and certain forms of cancer, it would provide benefit to consumers (better health). But it would also restrict their access to something they may genuinely want (a juicy steak).

In this paper, I am primarily interested in this more narrow kind of paternalistic regulation, which is associated with an actual tradeoff of benefits for individual consumers. I do not want to speculate which proportion of data protection regulations are of this kind. However, given that many potentially risky practices involving user data (e.g., the storage and sharing of large collections of interconnected user data) are also crucial to the business model of many online service providers (data analytics and targeted advertising), it is likely that at least some regulation of these practices will also affect the range and shape of products and services on offer.

That said, regulatory interference of this kind is a commonly used device in liberal societies, from requiring people to wear seat belts while driving or riding

in a car to restrictions on the sale of legal drugs like alcohol and tobacco. In order to assess if data protection warrants this kind of interference, as a society we should aim to understand how such regulation compares, normatively, to other common instances of liberal paternalistic intervention. Practically, the question liberal societies are confronted with is the following: given the extent of regulatory interference we accept in a range of other cases, which degree of interference, if any, should we accept in the case of data protection regulation? In this section of the paper, my aim is to show that trying to locate data protection regulation in relation to other forms of regulation on a map of regulatory interference does not yield satisfactory results. In other words, I intend to show that data protection regulations easily fall through the cracks of our intuitions about the appropriateness of current regulatory practices.[4]

For this purpose, I will analyze regulation of data protection along two dimensions. First, the "restrictiveness" of regulation, by which I refer to the extent to which that regulation interferes with a person's individual freedom to act as they see fit. Second, the clarity and tangibility of the calculus of risks, harms, and benefits, *i.e.,* the extent to which it is clear, both objectively and to individuals, what they stand to gain from regulation and how likely it is that individuals will end up in a situation in which they benefit from the regulation. Both these dimensions are significant from a liberal perspective:

1. **restrictiveness:** less restrictive regulations are easier to justify because the individual person needs to give up less of their personal freedom in order to comply with them.
2. **clarity of risk, harm, benefit calculus:** the more obvious and tangible the benefits of a restrictive measure, the higher the likelihood that the restrictions are congruent with reasons that a person subject to those restrictions recognizes themselves to have for restricting their freedom out of their own volition.[5][6]

Considering data protection regulation along these two dimensions will help us sketch a mental map of such restrictions from a liberal perspective and provide us with an idea of how this kind of regulation is relevantly similar or dissimilar to other kinds of paternalistic restrictions.

[4] In section 4, I will discuss why that is the case.

[5] The liberal concern for providing acceptable reasons echoes – though in weaker form – the considerations underlying the idea of consent: that the transfer of authority must be based on, or at least allude, in some way to individual volition. This idea runs through liberal theory from J.S. Mill ("Over himself, over his own body and mind, the individual is sovereign") to more recent debates about what shape this acceptability requirement should take in the context of liberal legitimacy. (See, Mill 2006, 16; Rawls 2005, 137; Waldron 1987, 36-37.

[6] This is not to say that there is no independent harm in losing my freedom to act one way or the other, even if I would not have taken advantage of it anyway and would have acted in line with the regulations because I thought it to be the right course of action all things considered, rather than for fear of punishment. But that's beside the point here. I do not mean to argue that there is definitely *no* loss in externally restricting individual freedom. I merely claim that being compelled to act in line with my own judgment is less bad than to be compelled to act on the basis of justifications which I do not find compelling.

		clarity of risk, harm, benefit calculus	
		clear	*unclear*
restrictiveness	*low*	seat belt laws	bike helmet laws
	high	alcohol and tobacco regulations	?

Fig. 1 examples of paternalistic regulation in comparison

In the remainder of this section, I will consider examples that fall within three different categories of paternalistic interference as determined by the interplay of the two dimensions introduced above, restrictiveness on the one hand and clarity and tangibility of risks, harms, and benefits on the other (see figure 1):

- seat belt laws have low restrictiveness, and their risks, harms, and benefits are clear and tangible;
- bike helmet laws have low restrictiveness; the risks, harms, and benefits are unclear;
- tobacco or alcohol regulations have high restrictiveness; but the risks, harms, and benefits are clear and tangible.

These different categories of intervention differ a great deal in terms of how widespread and how well accepted they are. Seat belt laws are both widespread and widely accepted, while regulations requiring cyclists to wear helmets are both rarer and more contested, and interventions that restrict smoking or alcohol consumption are common, but nevertheless often controversial. I will argue that the regulation of data protection does not neatly fit any of these categories, with both its restrictiveness and the calculus of risks, harms, and benefits remaining unclear.

3.1 Don't buckle up your data: data protection regulations are unlike seat belt laws

Let's start our analysis in the upper left corner of figure 1. Let's look at the first step in justifying seat belt laws:[7] there is clear scientific evidence demonstrating the effectiveness of seat belts at reducing the risk of death and injury for individuals wearing seat belts in traffic accidents, and this scientific consensus is largely mirrored in the public consciousness. People broadly agree that not wearing a seat belt greatly increases the risk of harm to them should they end up in an accident.

Now, consider the following justification for data protection regulation: the potential harms of data breaches are great and the benefits of regulating

[7] I assume here that seat belt laws are mainly justified on the basis of their benefits for the individual wearing the seat belt. Wearing a seat belt may also protect others, for example by preventing passengers from being catapulted out of the car in an accident and thus creating additional hazards for other road users who had hitherto not been involved in the accident. I take it that these other-regarding benefits of mandatory seat belt-wearing are at least not the sole objective of seat belt laws and that these laws are essentially paternalistic in character.

the use and transfer of data are obvious and will be widely accepted once people have been educated about them. At first glance, the case for greater general regulation of data protection looks a bit like the case for seat belt laws. Take the risk of identity theft, for example. Let's assume that, over time, I provide a range of personal information to various digital service providers: among other things, I provide my address, phone number and credit card data to an uncounted number of online shops. I mention my date of birth, full educational history and current employer, as well as a variety of other personal data on Facebook. I end up divulging my cat's middle initial and my mother's maiden name in response to funny Twitter memes, and hand over my social security number to the ever-so-useful online service which takes the pain out of preparing my tax return. I also leave a trail of behavioral information by way of cookies and other trackers embedded in the websites I visit, giving service providers information about which other online services (and even offline services if Bluetooth beacons are used) I use or engage with. Much of this data does not remain with the initial service provider. It is sold on for marketing and analytical purposes and may thus be disseminated far beyond the number of providers I used directly. In the (not uncommon) case of a data breach at one of these service providers, malicious actors collecting that data have access to enough personal information to, for example, apply for a mortgage on my behalf.

This harm is arguably not quite as bad as dying in a car accident but the harm is still quite substantial. Fighting identity theft and proving that it really was not me who applied for the mortgage can result in protracted and costly legal proceedings and, at minimum, causes hassle. Of course, there are other instances of identity theft and the severity of the associated harms varies. Credit card fraud, for example, is a fairly unsophisticated form of identity theft and has become a common occurrence, but is mostly a nuisance due to banks' technical abilities to distinguish fraudulent transactions from genuine ones.[8] Those who want to make the case for treating data protection regulation like seat belt laws can point towards this range of harms and risks and argue that they are so obvious and tangible that there is – or can be, once people have been appropriately informed – broad consensus on the benefits of restrictions on the storage, transfer, and processing of customer data.

However, while these harms may be obvious, the calculus of harms and benefits is more complicated in the case of regulating data protection than it is in the case of asking people to wear seat belts. The consequence of having to wear a seat belt is that I wear a seat belt. One possible consequence of tightly restricting to which extent Twitter (as just one example of an online service provider) can sell the data they collect on my interests, social group membership, and beliefs is to destroy their current advertising-based business model. It is by no means clear that this is a definite consequence of this kind of regulation, but it is not far-fetched to assume that serious regulation affecting

[8] Which, for instance, is a positive effect of the vast amount of data that is available to them which allows them to develop sufficiently detailed models of behaviour for their customers in order to identify deviant patterns.

their primary source of income might require them to change the way in which they monetize their services, such as switching to a subscription-based model. For my point to hold, it is not important whether this is what will happen. What is important is that there is uncertainty about what the landscape of data-reliant online services will look like once they are subject to stringent data protection regulations.

I stipulated earlier that regulation can be justified more easily if its calculus of risk, harm, and benefits is obvious and broadly aligned with individuals' personal preferences. In order for that to be the case, we must be able to discern what the likely consequences are. At the same time, this calculus also requires individuals to have an idea of the risks they face without regulation. However, most individuals tend to have at best an abstract idea of how giving up more data might be harmful to them. They might be aware that personalized data profiles could be used for purposes ranging from manipulating their choices or opinions to identity theft, but most users do not understand if and how, for example, the specific information an online news website collects about their interests and behaviour on their site and beyond contributes to those threats. If we cannot tell with sufficient confidence what might happen both as a consequence of regulation and absent any regulatory interference, individuals are not in a good position to assess if either is a scenario they find compelling. If I do not know if as the result of regulation, the online newspapers I currently read for free will move to a subscription model and if there will still be a provider who hosts my blog for free, it is hard for me to figure out what option – regulation vs. no regulation, or what degree of regulation – is best aligned with my interests. In other words, without a grasp of the consequences of regulation, we face a problem on the epistemic dimension of my framework (we are moving to the right on that dimension in figure 1). I don't intend to judge whether that deficiency is enough to defeat regulatory efforts. For my purposes in this paper, it is enough to conclude that it weakens the basis for justifying paternalistic regulation.

This is consistent with what we can observe in the debate on making bike helmets mandatory for adults. This case is very similar to making it mandatory to wear a seat belt while in a car when it comes to the restrictiveness of the measure – leaving aesthetic considerations aside – except that the evidence for the effectiveness of helmet laws is mixed. While the evidence is clear that wearing a helmet while cycling protects the individual from serious head injury, (Attewell et al. 2001) population-level studies do not find a reduction in the incidence of such injuries among cyclists in countries which have made the wearing of helmets mandatory (Robinson 2006). The individual trying to figure out if the demands of a mandatory helmet law are aligned with her own judgment is left with conflicting evidence. While wearing a helmet should make her safer, the fact that she lives in a society where doing so is mandatory may not. Hence, proponents of helmet laws do not have a clear-cut case that being required to wear a helmet is actually aligned with my preferences (assuming, for the sake of the argument, that increasing my safety is what I value most on this occasion).

I do not want to overstate the similarity between data protection regulations and helmet laws. My main purpose was to illustrate roughly the space of issues we are entering when we move along on the epistemic dimension of the framework. Helmet laws are comparatively rare.[9] I am in no position to judge whether this is actually in part because of the difficulty to demonstrate conclusively that they further individual cyclists' preferences for their own safety. But the absence of a *clear-cut* justificatory case for increases in cyclists' safety is certainly not helpful.

In summary, helmet laws differ significantly from seat belt laws in that they lack a clear and tangible calculus of the risks, harms, and benefits associated with the space they meant to regulate. If anything, this calculus is even more complex and thus more opaque to most individuals in the case of regulating data protection.

3.2 Call off that data sale? Restrictiveness of data protection regulations

So far, we have established that the justificatory basis for paternalistic data protection regulation is weakened by comparison with our initial point of comparison – seat belt laws – because the risk, harm, and benefit calculus is not clearly in favor of such regulation. So how does it fare on the other dimension – restrictiveness? The case for the mandatory wearing of seat belts is not just supported by the fact that it provides obvious and uncontested benefits. Making people wear a seat belt is also a fairly small infringement on their individual freedom. Wearing a seat belt does not substantively alter the way in which people use their car. They can still drive to the same places, at the same speed, whenever they determine they want to go. It is just that the seat belt needs to be worn while carrying out these activities. (This is not to say that being required to wear a seat belt does not infringe upon their personal freedom at all. If they have an aesthetic preference for driving without a seat belt, or a nostalgic longing for driving along the French Riviera like they do in a cheesy 1950s movie, they will not be able to act on these preferences.)

I am not interested in determining exactly where seat belt laws are located on a scale of the restrictiveness of coercive regulations or whether any restrictive measures, however small, can ever be justified for paternalistic reasons. If they cannot, there is no need to consider if data protection meets the threshold to be within the purview of paternalistic regulation. But insofar as we do accept some degree of paternalistic regulation, it makes sense to think about where on the (multi-dimensional) continuum data protection regulations fall.

So does regulating online service providers' capacity to collect, store, and transfer data restrict their *users* – *i.e.,* the freedom of those who are supposed to benefit from this paternalistic intervention – personal freedom? At first glance, the obvious burden of such regulation is imposed upon the service provider who will have to adapt their business practices in order to comply

[9] Only three countries (Argentina, Australia, and New Zealand) currently have universal mandatory helmet laws.

with the regulation. Depending on the scope and specific aspects of the regulation, we can imagine effects ranging from "merely" requiring providers to adapt their technical infrastructure (*e.g.,* change the technical design and implementation of their database in order to comply with users' right to have all their information erased[10]) to the need to revise their business model, such as moving from an advertising-based to a subscription-based model, as I discussed earlier.

However, the situation is more complicated. Indirect restrictive effects of tightly regulating what providers can do with their users' data can also hit those users. For example, regulation that interferes with companies ability to collect data about their users or to monetize that data,[11] indirectly also restricts the user's capacity to enter a contract that allows them to effectively use their data as currency. By shaping the kind of Internet that is economically viable, regulation also shapes the types of exchanges that are possible. Curtailing the opportunity for individual users to hand over their data in exchange for services may not necessarily be the primary target of the regulation but if it renders this kind of exchange economically unviable, their freedom to "sell" their data is *effectively*, though not legally, curtailed. So, compared to making people wear a seat belt, interfering with online service providers' collection and use of customer data may be less direct in restricting the way in which individuals can act, but to the extent that it ends up having an indirect effect, it substantively diminishes their *actual capacity* to engage in an economic activity of a specific kind between mutually consenting parties. This is not to say that such regulation cannot possibly be justified, or that individuals have an antecedent right to this particular kind of relation, but only that it is far from neutral with respect to individual users' economic freedom. In fact, the extent to which such regulation could infringe upon individuals' personal freedom is – depending on the specific content and scope of the regulation – potentially quite severe, since it can disable or significantly curtail a kind of activity in its entirety (individuals using their data as currency), instead of interfering merely with the way in which this activity is executed (such as driving *with* a seat belt).

Paternalistic regulation which significantly infringes upon individuals' personal freedom is not uncommon across liberal societies. For example, the distribution and consumption of certain recreational drugs such as tobacco and alcohol is generally highly regulated – though to varying degrees – in liberal societies across the board. But even those regulations fall short of entirely disabling the activity in question – *i.e.,* the actual consumption of the drug

[10] This provision, for example, is part of the European Union's General Data Protection Regulation (GDPR), the California Consumer Privacy Act (CCPA), and the proposed U.S. Consumer Online Privacy Act (COPRA).

[11] This is how most online services which do not rely on a subscription model or are backed by sponsors work. Now, many people are not aware that this is how those services work and that they are exchanging their data for services. The fact that people may be unaware that they are effectively paying with their data is a good argument for increasing transparency about that fact, but not for preventing people from entering into such an exchange if they so wish.

		clarity of risk, harm, benefit calculus	
		clear	*unclear*
restrictiveness	*low*	seat belt laws	bike helmet laws
	high	alcohol and tobacco regulations	data protection regulations

Fig. 2 data protection regulations in context of other paternalistic regulations

– despite the fact that there is clear evidence that common and widely accepted levels of consumption carry significant health risks (Kreitman 1986). That said, the consumption of a majority of hard drugs is actually prohibited. The rationales behind the regulation of drugs is far from consistent and the history of the regulation of drugs is complex. However, alcohol and tobacco are still interesting cases because they indicate that the burdens for actually paternalistically disabling a particular kind of activity which is currently legal, commonly practiced, and widespread throughout many strata of society are quite high. Liberal societies tend to not fully disable widely accepted risky behaviors even where the calculus of risk, harm, and benefits is clear and tangible.

This makes sense from a liberal perspective. In cases where the cost of interference is high, the weight of the objective, external calculus takes a back seat to the prerogative of autonomous self-determination. In other words, when the stakes are high, – potentially disabling, rather than merely modifying or adding to an activity – the liberal calculus skews towards allowing individuals to assess the risks, harms, and benefits of an activity in light of their own preferences and objectives. They are allowed to discount future harms and benefits, to act on the basis of biases which are specific to them, and in general to revise their decision based on factors that fail to be captured by the objective calculus. Given this constraint, the liberal bar for fully disabling an activity on purely paternalistic grounds, that is, for the protection of the person pursuing it, is extremely high.

So, where on the restrictiveness dimension can we locate data protection regulations? I concluded earlier that – depending on the specifics of the regulation – they hold the potential to actually disable a particular kind of activity (individuals using their data as currency). Curtailing individuals' ability to sell their own data may not be the explicit target of such regulations, but insofar as the disabling effect is a clear consequence of the regulations, their effects seem to be potentially more restrictive than some of the most intrusive kinds of paternalistic interference currently commonly tolerated in liberal societies.

Paternalistic data protection regulations thus face challenges on both of the dimensions I considered in this paper: they potentially score very high on the scale of the restrictiveness of the interference. At the same time, they lack support by a clear and conclusive objective calculus of risks, harms, and benefits. To the extent that these regulations affect them, individuals face the prospect of being curtailed in their ability to engage in particular kinds of transactions, but left without a clear and conclusive argument as to why they should think that the effects of the interference will be overall beneficial to them. This means that determining where data protection regulations fit in

the context of existing regulatory interventions is hard: they are as restrictive or more restrictive than some of the most restrictive paternalistic regulations found in liberal societies, without being supported by an tangible calculus of risks, harms, and benefits which is clear and uncontroversial.

In other words, at least some forms of data protection regulation (*i.e.,* those with high restrictive potential) appear to be outside the scope of current regulatory practices and require us to take a more in-depth look at the specific reasons that support particular instances of regulation. However, any effort in that regard is seriously hampered as long as individuals are unable to clearly assess the calculus of risks, harms, and benefits of the data economy and data protection. In the next and final section, I will turn to explaining why liberal societies – both on an individual and on a collective level – struggle to make these calculations and argue that some regulatory efforts can be justified for the sake of enabling us to make these assessments.

4 Digital Bodies

Our difficulties to determine what kind and degree of data protection regulation of online service providers is appropriate, as well as our intuitive, but ultimately futile reliance on consent have the same root cause: the relationship between the individual and their environment has changed fundamentally where those interactions have shifted online in ways that make it hard for individual persons to perceive, reason about, and control them. Each and every of those interactions consists of exchanges of data, some of which we provide consciously (such as when we enter our name and address in the process of ordering from an online store), some of which is generated as a by-product of our actions online: data of that kind is a record our behaviour (e.g, how much time I spent looking at different items), preferences (*e.g.,* based on the other kinds of items I considered), and connections (*e.g.,* if I access the online store via a recommendation on social media). We move around the online world enveloped in a cloud of data which is an imprint of ourselves and our actions online. Others (*e.g.,* providers of digital services) can access, analyse, and transfer this data in order to learn more about us and to affect our behaviour (for example through targeted advertising).

In other words, by using a variety of online services which are linked by tracking devices or central ownership of different platforms, we generate a *digital body* in addition to our physical one, which can be observed and used to influence us in turn. With our physical bodies, we know when we can expect to be observed and have a sense – though not necessarily a perfect one – of what kind of information others might derive from observing us and how it is likely to be used. When I walk to a shop in order to buy a newspaper, I know that my neighbours might see me go out, that I run the risk of getting robbed on the way if I walk there late in the evening, that the shop is likely to use surveillance cameras to detect shoplifters, and that the person behind the counter might recognize me, notice that I am buying the same paper as

usual and recommend a magazine I might like. If we feel uncomfortable with any of this, we also have at least some idea of how to obscure our physical bodies in order to avoid observation and consequent attempts at influencing our behaviour. I might use the backdoor if I do not want my neighbour to see me go out, walk to the shop during the day in order to avoid getting robbed, wear sunglasses and a hat in order to stay incognito at the store, or patronise different stores so the employees do not recognise me as a regular.

Compared to these intuitions about how to evade observation or obscure the presence of our physical bodies, out intuitions about the risks and harms of acting in the digital sphere, let alone on how to avoid them are fairly poor, as I have argued earlier.

The following three characteristics go some way to accounting for this blunting of our intuitions when it comes to our digital bodies. Our digital bodies are:

1. easily observable and ubiquitously observed by others by default,
2. intangible to ourselves, and
3. interconnected.

They are *easily observable and ubiquitously observed* because all action online essentially consists of transfers of data. The threshold for others to record and analyze that data is low and is often necessary for providing the service in the first place. Using and storing data for purposes beyond providing strictly necessary functionality is easy and the line between what is necessary and what is optional is easily blurred.[12] As a result, we have to make decisions about our data all the time – when deciding whether to sign up for that newsletter in order to get the 10% discount next time, or when figuring out whether we would rather keep on our adblocker and anti-tracking browser extension or be able to read that interesting article a friend just sent us – but the sheer amount of data which we are expected to consider in each of those instances make it hard to substantively reflect on each of those decisions.

This kind of cognitive overload is exacerbated by the fact that our digital bodies are also *intangible* to us. Even people who are generally aware of the fact that their data is constantly collected and analysed mostly lack a sense of what information precisely is revealed during which activity online, and how it might be used both in the short and in the long term. This is akin to being unable to perceive and relate to the left hand of our physical bodies and only noticing that we are hurt after cutting our finger with a kitchen knife when we are about to pass out from the blood loss. If we do not even perceive our digital bodies and what they reveal about us, it is hard to connect specific actions with potential threats. For example, it is hard for me to tell which of my previous actions (and the data I provided by pursuing it) made me a target for a political misinformation campaign if I am unaware of the majority of data which has been generated about me over the course of my life online so far, let alone who may have received access to that data in the meantime.

[12] See, for example, considerable ambiguity over what constitutes a "legitimate interest" for data processing under the GDPR (U.K. Information Commissioner's Office (ICO) 2020).

In other words, our senses fail us because our digital bodies in their current state are essentially numb.

Importantly, that numbness does not just have an impact on each of us individually, it also affects others since our digital bodies are *interconnected.* Personal identifying information is often part of the data that online service providers collect and store (be it information provided by us or the unique data fingerprints derived from our device and location settings) and is easily linked to that of other individuals who we interact with online, *e.g.,* by engaging with friends or people who share our interests on social media or by sending a "friends and family" discount coupon from my favourite online shop to my brother. Data profiles which can be linked in this way allow for inferences to be drawn from data about ourselves (our preferences, characteristic patterns of activity) to personal information, traits, and likely behaviours of those who are in some way connected to us. For example, becoming a parent and buying baby accessories might increase the odds that individuals connected to me who are in a similar demographic might receive baby-related targeted advertising or political campaign materials directed specifically at young families based on the prediction that they will or might soon react favourably to it as well. Similarly, though seemingly less innocuous, my father's donation of his DNA to a genealogical database in order to find new relatives and extend the branches of his family tree might draw me into the focus of law enforcement who use such DNA to generate new leads on cold cases. Because our digital bodies are interconnected in this way, the potential vulnerabilities which result from the data we provide about ourselves also extend to those we are connected with. Our digital bodies are not just a danger to ourselves but also to others, but given their intangible nature, we are mostly unaware of that danger.

In conjunction, these three characteristics of our digital bodies – ubiquitous observability, intangibility, and interconnectedness – do their part to impede our understanding of our own vulnerabilities but also of the danger we pose to others. Without such an understanding, it is clear that individuals online lack agency. That is, they lack the capacity to meaningfully reflect on the risks, harms, and benefits, draw conclusions on the trade-offs acceptable to them both as individuals and collectively, and to exert control over the way in which their digital bodies interact with their environment. If we cannot grasp the full scope and potential implications of our actions online, our capacity, as agents, to act in accordance with our values and convictions about how we want our lives to go is seriously diminished.

Individually, being aware of the vulnerabilities of our digital bodies is crucial to developing a sense for which interactions of our digital bodies with those of others and with online service providers might expose us to levels of potential harm and how to effectively avoid or defuse those interactions. As contributors to such a debate, the task for political philosophers, and specifically for liberal theorists, is to address what degree of data control and protection is required to develop and preserve individual autonomous agency online. Equally, computer scientists and engineers are faced with the question what kinds of technical solutions are required in order to achieve those goals.

One conclusion that we can draw even prior to engaging in those debates is that if we are unable to see our digital bodies clearly, we are definitely deprived of *political* agency. If, as a society, we cannot pinpoint the sources and extent of our vulnerabilities in the digital sphere, we are likely to continue stabbing in the dark when we try to determine the necessary and appropriate level of regulation or fail to acknowledge the loss of effectiveness of the classic liberal paradigm of consent. On a political level, this awareness of the nature of our digital bodies is crucial for generating an effective public debate and developing a collective sense for which of the potential harms we face as a result of our new data-based representations online warrant what kind of regulation. In other words, if we want to take any steps to protect individual autonomous agency online, we must take steps to make digital bodies visible and relatable for the average person.

In the first place, this justifies forms of regulation that make transparent what data concretely online service providers have collected about us, how they process it and who else has access to it. This form of regulation is already gaining traction. The GDPR, for example, grants users the right to request all data an online company has stored about them (European Union 2016, Article 15). However, this transparency regulation is only effective for improving individuals' understanding of their digital bodies if the amount of the information provided is not so great as to be overwhelming. Hence, measures to increase data transparency should be accompanied by regulation that limits the amount and complexity of the purposes for which data can be used, and imposes limits on the longevity of personal information in the databases of online service providers (*e.g.,* deletion of accounts by default after some time of inactivity). Contextual information in *accessible* language about the purposes and risks of storing and processing that data (akin to patient information sheets included with pharmaceuticals) would also help to make our digital bodies more tangible to us. This is not a comprehensive list of measures and merely sketches the direction for a first layer data protection regulation if we look at it through the prism of digital bodies. But the gist is clear: we must adjust the size and complexity of the imprints of data by which we are represented in the digital sphere – our digital bodies – to our limited mental processing capacity, so we can actually perceive and reason about them in a meaningful way.

5 Conclusion

At a time when increasing parts of our daily lives are shifting online, liberal societies need to sharpen their gaze for how this shift has also transformed the character of our interactions in the digital sphere. The upshot of the ideas I presented in this article is twofold:

1. the idea of the digital body provides a prism for assessing those interactions and explains why classic liberal paradigms of consent and regulation fall short so far: I argued that the ubiquitous observability, intangibility, and

interconnected nature of our digital bodies is too overwhelming for the individual to allow for meaningful decisions when asked for consent. These same characteristics also make it hard for individuals to clearly assess the severity of the potential risks and harms associated with their use of data-driven services online and balance them with the benefits; and
2. considering questions of data protection with the idea of the digital body in mind gives us an idea of what first steps we must take to approach this impasse and start a productive debate about the prerequisites of agency online.

All of this means, that, somewhat paradoxically, in order to even start an effective public debate on how to react to the challenges of the digital realm as liberal societies, we must first enact some regulation that creates transparency and generates concrete awareness about the imprint of personal data we leave with every interaction online. We must hand individuals the tools for engaging in that debate and enables them to reason clearly about their own vulnerabilities and potential for harming others. In other words, we must show ourselves our digital bodies.

References

Attewell R, Glase K, McFadden M (2001) Bicycle helmet efficacy: a meta-analysis. Accident Analysis & Prevention 33(3):345–352, DOI 10.1016/S0001-4575(00)00048-8, URL http://www.sciencedirect.com/science/article/pii/S0001457500000488

Blum D (2018) The Poison Squad: One Chemist's Single-Minded Crusade for Food Safety at the Turn of the Twentieth Century. Penguin, New York

European Union (2016) 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). Official Journal of the European Union L119:1–88, URL http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC

Kitkowska A, Warner M, Shulman Y, Wästlund E, Martucci LA (2020) Enhancing privacy through the visual design of privacy notices: Exploring the interplay of curiosity, control and affect. In: Proceedings of the 16$^{\text{th}}$ Symposium on Usable Privacy and Security (SOUPS), pp 437–456, URL https://www.usenix.org/conference/soups2020/presentation/kitkowska

Kreitman N (1986) Alcohol consumption and the preventive paradox. British Journal of Addiction 81(3):353–363, DOI 10.1111/j.1360-0443.1986.tb00342.x, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1360-0443.1986.tb00342.x

Locke J (1988) Locke: Two Treatises of Government. Cambridge University Press, Cambridge

Mill J (2006) On Liberty and the Subjection of Women. Penguin, London

Rawls J (1972) A Theory of Justice. Oxford University Press, Oxford
Rawls J (2005) Political Liberalism, 2nd edn. Columbia University Press, New York
Raz J (1986) The Morality of Freedom. Oxford University Press, Oxford
Robinson DL (2006) No clear evidence from countries that have enforced the wearing of helmets. British Medical Journal 332(7543):722–725
Shuster E (1997) Fifty years later: The significance of the nuremberg code. New England Journal of Medicine 337(20):1436–1440
Tsai L, Wijesekera P, Reardon J, Reyes I, Egelman S, Wagner D, Good N, Chen JW (2017) Turtle guard: Helping android users apply contextual privacy preferences. In: Proceedings of the 13th Symposium on Usable Privacy and Security (SOUPS), Santa Clara, California, USA, pp 145–162, URL https://www.usenix.org/conference/soups2017/technical-sessions/presentation/tsai
UK Information Commissioner's Office (ICO) (2020) When can we rely on legitimate interests? URL https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/legitimate-interests/when-can-we-rely-on-legitimate-interests/
Waldron J (1987) Theoretical foundations of liberalism. The Philosophical Quarterly 37(147):127–150
Waldron J (1993) Liberal rights: collected papers, 1981–1991. Cambridge University Press, Cambridge
Wei M, Stamos M, Veys S, Reitinger N, Goodman J, Herman M, Filipczuk D, Weinshel B, Mazurek ML, Ur B (2020) What twitter knows: Characterizing ad targeting practices, user perceptions, and ad explanations through users' own twitter data. In: Proceedings of the 29th USENIX Security Symposium (USENIX Security), pp 145–162, URL https://www.usenix.org/conference/usenixsecurity20/presentation/wei