AI & SOCIETY
https://doi.org/10.1007/s00146-023-01734-w

OPEN FORUM

The privacy dependency thesis and self-defense

Lauritz Aastrup Munch[1] · **Jakob Thrane Mainz**[1]

Received: 5 December 2022 / Accepted: 18 July 2023

Abstract
If *I* decide to disclose information about myself, this act may undermine *other* people's ability to conceal information about them. Such dependencies are called privacy dependencies in the literature. Some say that privacy dependencies generate moral duties to avoid sharing information about oneself. If true, we argue, then it is sometimes justified for others to impose harm on the person sharing information to prevent them from doing so. In this paper, we first show how such conclusions arise. Next, we show that the existence of such a dependency between the moral significance you are inclined to attribute to privacy dependencies and judgments about permissible self-defense puts pressure on at least some ways of spelling out the idea that privacy dependencies ought to constrain our data-sharing conduct.

Keywords Privacy rights · Privacy dependencies · Self-defense · Surveillance · Genetic tests

1 Introduction

Consider the following three cases:

Identical twins. Smith and Jones are monozygotic twins. Nora is a medical doctor and she knows that Smith and Jones have high genetic similarity. Smith decides to share his DNA profile with Nora so that she can check if he has a predisposition to develop certain diseases. Since monozygotic twins share their DNA profile, any information Nora retrieves from Smith's DNA profile is also information that Nora now has about Jones. Jones, however, would not like Nora to know what diseases he is predisposed to develop.

Bank loan. Peter and Tom are customers of the same bank. Peter applies for a loan. To decide if Peter is eligible, the bank decides to train an advanced machine learning model that can estimate—based on all the information the bank has on Peter—whether he will default on the loan. The bank makes Peter an offer: if Peter is willing to give up huge amounts of information about himself, he will receive a loan with much lower interest rates, *if* it is decided that he is eligible. Peter accepts the offer. As it turns out, the machine learning model was not very accurate, and it categorized Peter as eligible for the loan, even though he eventually ended up defaulting on the loan. Now Tom applies for a similar loan. Having learned from their mistake, the bank has now updated the model so that it more accurately predicts loan eligibility. Peter and Tom are very similar in many ways. They are roughly the same age, they live in the same neighborhood, they have the same credit history, the same income, and so on. But because the model has been updated, it now predicts that Tom is *not* eligible for the loan and the bank therefore turns down his application. Tom would have preferred that the bank did not predict his eligibility.

Political affiliation. Ben and Tim are partners. They live together, and they both use the computer they have in the living room. Ben fully endorses political party X. Tim, on the other hand, is a strong opponent of X, and he much rather prefers the competing party Y. Party X and party Y have hired the same data analytics company to help them target voters with political ads to convince them to vote for their respective parties. The data analytics company has a hard time finding out who is using the computer in the living room, because the search histories of Ben and Tim are so different. However, the company finds out that Ben shares on his social media profile that he fully endorses political party X. Since the company knows that only Ben and Tim use the computer in the living room, they infer that Tim must be using it when Ben is not. So, the company decides to target Ben with manipulative ads for party Y when he is using the

✉ Lauritz Aastrup Munch
laumu@cas.au.dk

Jakob Thrane Mainz
Jakob-mainz@hotmail.com

[1] Department of Philosophy and History of Ideas, Aarhus University, Jens Chr. Skous vej 7, Aarhus, Denmark

Published online: 19 August 2023

computer, and target Tim with manipulative ads for party X when he is using the computer. Tim really hates political ads, and he would have preferred that the company did not know his political affiliation.

These three cases are structurally analogous. Their common structure is this: there is an agent, A, who share their information *p* with another agent C, and by doing so, C is put in a position where they can infer information *q* about B, a third agent, who would prefer that C did not infer *q* about them. Such dependencies are called *privacy dependencies* (Barocas and Levy 2019). Privacy dependencies raise intriguing normative questions. This is primarily because stock theories of the right to privacy say that people are owed some level of control over their information. But, as the existence of privacy dependencies show, I can sometimes undermine the privacy of others by sharing information about myself. So how should this fact constrain my conduct, if at all?

One possible response would be to say that *whenever* sharing one's information would undermine the kind of control over their privacy others are normally entitled to, this fact should tell decisively against me so doing. This seems to be the kind of view Carissa Véliz alludes to in her recent book:

> "Our interdependence in matters of privacy [privacy dependencies, red.] implies that no individual has the moral authority to sell their data. We don't own personal data like we own property because our personal data contains the personal data of others. Your personal data is not only yours (...)" (Véliz 2020: 79).

What Véliz seems to be saying is this. When you share information about yourself, you may end up revealing private things about others. And given that we would normally say that other people have a right to privacy that entitles them to decide if these private things are revealed, we should think this right would continue to be applicable even in some cases where you reveal something about yourself that happens to reveal something about others.[1] In the relevant range of cases one will thus lack the moral authority to authorize the selling (or mere sharing) of one' own information, because it 'contains' the personal data of others, and thereby implicate their privacy. More generally, if we believe some people lack the moral authority to authorize the sharing of their information when there exist privacy dependencies, then these people would have a duty not to share it (unless the relevant authorization was given from some other source, presumably from those whose privacy is affected). Call this view the *Privacy Dependency Thesis* (PDT):

> Privacy dependency thesis (PDT): Agent A has a duty not to share information *p* about A with agent C, if doing so makes it possible for C to infer information *q* about agent B that B otherwise has a right to keep private.

As we shall see, the PDT or close variations of it, enjoys widespread support in the literature (e.g., Mühlhoff 2021; MacCarthy 2011; Véliz 2020). This might be surprising given its far-reaching implications. The thesis suggests that many of us are currently wronging others when undertaking activities that might at first sight seem innocuous. Many of us happily share our information (online as well as and offline) giving much thought to it may be used to reveal information about others: we create social media accounts, we share pictures of ourselves, we expose our opinions and commitments on online fora, and so on. If the PDT is true, then many of us are acting impermissibly in our seemingly innocuous daily online behavior.

In this paper, we argue that the PDT (at least in the formulation above) gives rise to an intriguing challenge: if the PDT is true, it follows that people whose privacy is negatively affected by other people sharing personal information about themselves have a permission to impose harms on those others as a means of preventing them from sharing personal information—as an act of justified self-defense. This is because PDT effectively picks out a sense in which one may be causing an unjustified threat to others—something that standard theories of permissible self-defense deem sufficient grounds for imposing harm on others (Lazar 2012). The interesting upshot, then, is that one should *either* accept this revisionary account of what people may do to protect their own privacy, *or*, if one deems this unacceptable, find oneself with a compelling reason to revise the scope of the PDT. Those that are deeply committed to the PDT will expectably be happy to learn more about what follows from their commitment. Others might find that the implication we derive constitutes an effective *reductio* on this particular version of the PDT.

The paper is structured as follows. In section II, we explain in more detail what a privacy dependency is, and we show how advocates of the PDT defend it. In section III, we outline our challenge to the PDT as formulated above. In section IV, we discuss and reject four objections. These say, respectively, that self-defense need not be justified after from the PDT because (i) those who share their information will not be liable; (ii) those that share their information will not make a causal contribution; (iii) in the relevant cases there are more apt targets of self-defense available, meaning that those who share their information are not permissible targets of self-defensive harm; and (iv) that privacy rights

[1] Notice that although Véliz' claim is about the impermissibility of selling data specifically, it's hard not to take her idea to generalize. This is because selling doesn't seem to be relevantly dissimilar from other ways in which one might waiver one's entitlements over information. If I may not *sell* my information (and associated use rights) to others, why should I be permitted to *give it away for free*?

are not enforceable. In section V, we offer a few concluding remarks.

2 The privacy dependency thesis

2.1 What is a privacy dependency?

Before taking a closer look at the PDT, let's start by saying more about privacy dependencies. According to Barocas and Levy, a privacy dependency concerns:

> "the varied ways in which one person's privacy is implicated by information others reveal. We term these phenomena *privacy dependencies* and we identify three broad types. In a tie-based dependency, an observer learns about one person by virtue of that person's social relationships with others—family, friends, or other associates. (...). In a similarity-based dependency, inferences about our unrevealed attributes are drawn from our similarities to others for whom that attribute is known. And in difference-based dependencies, revelations about ourselves demonstrate how we are different from others—by showing how we "break the mold" of normal behavior, showing how we rank relatively on some desirable attribute, or by allowing an observer to pinpoint an unknown person through process of elimination." (Barocas and Levy 2019, pp. 558–559).

Notice that Barocas and Levy speak of how privacy is implicated by "information others reveal". But that is too broad, at least if the notion of a privacy dependency is meant to pick out the narrower phenomenon described in the examples in the introduction. The reason is that there is a difference between, say, Alice "revealing information" about Bob, and—as Barocas and Levy put it—"that Alice may disclose information that is *explicitly and exclusively about Alice*, seemingly having nothing whatsoever to do with Bob, and can still implicate his privacy in so doing" (Barocas and Levy 2019, p. 559). We will be concerned with the latter, narrower, understanding of privacy dependencies. Why? Because there seems to be a moral difference between sharing other people's privacy-protected information, and sharing information that is "explicitly and exclusively" about you that can then be used to infer privacy rights-protected information about others. Whereas the former case seems obviously morally impermissible, it is at least less obvious that the latter case would be. Once we start looking for them, we see that privacy dependencies are pervasive. Some are obvious, like the one involving identical twins like Smith and Jones. Others are more subtle, like the one involving Peter and Tom in the bank loan case. These subtle dependencies have, however, become much less subtle by the relatively recent advent of machine learning and deep neural networks. With these novel data analytics technologies, it has become much easier to find patterns in huge datasets (Barocas and Levy 2019, p. 587). What this means is that we can now discover privacy dependencies that are prima facie invisible to the human eye. In the example of Peter and Tom, there is no causal link between Peter defaulting on the loan, and Tom defaulting on the loan. Even if one knows that people with the same demographics *tend* to have similar risks of defaulting, it need not be the case. The dependency is merely correlational. The PDT should be viewed against this background of such extraordinary new possibilities of discovering privacy dependencies. Since new technologies make it possible to infer so much information about people, we risk giving sensitive information away about other people, when we decide to share even trivial and seemingly innocent information about ourselves.

2.2 Motivating the PDT

Let us now have a look at how advocates of the PDT tend to motivate it. The PDT, or variations of it, has been defended by numerous philosophers and legal scholars alike.[2] Advocates of the PDT tend to produce either *rights-based* defenses or *harm-based* defenses of their view. More precisely, they tend to defend their view in the two following ways: they claim that (1) sharing information about yourself can—through the functions of privacy dependencies—constitute or facilitate privacy rights violations of others, and that (2) sharing information about yourself can—again through the functions of privacy dependencies—lead to serious harms of other people, for example by imposing financial burdens on them that set back their interests significantly. Some advocates of the PDT give *both* rights-based and harm-based defenses of their view, and it is not always clear to which kind a particular defense belongs. This is not a problem here, though, since we are merely invoking the distinction for presentational purposes. Perhaps the most notable defense of the PDT, which encompasses both types of defenses, comes from Carissa Véliz. Consider what she says immediately after the quote we saw above:

> "Our interdependence in matters of privacy [privacy dependencies] implies that no individual has the moral authority to sell their data. We don't own personal data like we own property because our personal data contains the personal data of others. Your personal data is not only yours. (…) your privacy slips can facilitate violations of the right to privacy of other people." (Véliz 2020: 79)

[2] Discussions of the PDT and variations of it can be found in Susser (2019), Fairfeld and Engel (2015), Reidenberg et al. (2014), Zarsky (2004), Véliz (2020), MacCarthy (2011) and Mühlhoff (2021).

With respect to (1), Véliz says that when you share information about yourself, you may facilitate the violations of the *rights* of other people—in particular their privacy rights. In the example of Smith and Jones, Véliz would probably say that Smith facilitates the violation of Jones' right to privacy by sharing his DNA profile with Nora. It is not clear what exactly Véliz and other advocates of the PDT mean by 'facilitation'. And admittedly, it does not always seem wrongful to facilitate rights-violations. But since we are interested in raising another challenge to the PDT, we will not delve into these details here.

Another defense of something in the close vicinity of the PDT comes from Rainer Mühlhoff. Although Mühlhoff's main goal is not to defend the PDT, he writes the following in passing:

> "The collective of data donors goes from being rights-holders to duty-bearers. *Predictive privacy makes it a duty for all of us, both in our roles as users and citizens, to ensure that no detrimental treatment of others is facilitated through the data* (including de-identified data and usage data) that we submit to platforms and digital services." [emphasis added] (Mühlhoff 2021: 680).

Mühlhoff stresses a defense like (2), by pointing to the *detrimental treatment of others* that can emerge because of privacy dependencies, and claims that, as a consequence of such possibilities, we have *duties* to ensure that this treatment does not occur. It is worth noticing that Mühlhoff also counts 'de-identification' as detrimental treatment, suggesting that he also has something like (1) in mind. On Mühlhoff's view then, *if* the information we share is likely to be used to bring about these detrimental effects, including privacy rights violations, then we have a duty not to share the information—which is essentially what the PDT holds.

Yet another example of a defense of something in the vicinity of the PDT comes from Mark MacCarthy. He writes:

> "Even when individuals have the ability to refuse data collection requests, if enough other people go along with the information collection and use scheme, the economic damage is done. An unfairness framework for privacy needs to supplement the informed consent model. *If the harm done by negative privacy externalities is substantial, then individual choice might have to be restricted*. Simply getting informed consent would not make an information practice legitimate." [emphasis added] (MacCarthy 2011, pp. 5–6).

While MacCarthy seems to think—like Véliz—that privacy dependencies can result in privacy *rights* violations, the quote above focuses more narrowly on the harmful *consequences* of these dependencies. Thus, MacCarthy primarily bases his defense on something like (2). And although MacCarthy phrases things cautiously, he seems to regard it as a real possibility that people's choices 'ought to be restricted' to prevent harm from data sharing. This suggests that people sometimes should not share data about themselves, because, plausibly, only if they should not there would be a basis for preventing them from doing it.

Notice that the harms alluded to above need not always be present in cases where information is revealed, and the harms, when they occur may take a variety of forms and differ in the severity. In the most egregious cases, they could be very significant. For instance, if insurance companies get their hands on your genetic profile, they might charge you more. In other cases, they might be mere non-trivial nuisances, where the main problem might be that it simply ought to be up to you to decide when the relevant information is disclosed.

Before we proceed, let's clarify how these considerations feed into the gloss we give on the PDT:

> Privacy dependency thesis (PDT): Agent A has a duty not to share information p about A with agent C, if doing so makes it possible for C to infer information q about agent B that B otherwise has a right to keep private.

Specifically, the rights-based defense and the harm-based defense might be thought of two separable ways of substantiating the latter requirement in the PDT, the claim that "B otherwise a right to keep private". It's easiest to see how the sentence is compatible with the rights-based defense. In cases where sharing information about oneself and this, via privacy dependencies, facilitate the violation of privacy rights, then it is true that A helps to reveal something that "B otherwise has a right to keep private" because A has a right to privacy. But although it can be read in this way, the formulation "has a right to keep private" need not refer to a 'right to privacy' specifically. It could also be given a reductive reading where the underlying concern, and reason why B has a right that something is kept 'private' (in the sense of hidden or undisclosed) has to do with avoiding harm. On this reading, whenever a sufficiently grave harm might ensue because of sharing one's information—and we are not saying that such harms will always or even mostly ensue; after all, the relationship between loosing privacy and harm seems contingent at best—then this satisfies the requirement that B has a right to keep the information private.

Finally, it is helpful to contextualize the PDT. It can be seen as part of a bigger literature that has, over the last decade, pressed serious criticisms of the paradigm way information transactions are regulated online; the 'notice-and-consent' scheme.[3] Against this background, the PDT specifies one reason why notice-and-consent might be inappropriate

[3] See Susser (2019) for an overview of this literature.

as a way of authorizing information transactions. More specifically, and nicely illustrated by Véliz' quote above, the PDT does so by pointing to a sense in which our data sharing choices have consequences for third-parties which therefore makes consent unable to morally authorize the transaction.[4] In this sense, the PDT echoes an influential line of reasoning in the privacy literature pointing out that privacy has 'social' and not only 'individual' value (Cohen 2012; Susser 2019; Nissenbaum 2010). However, many believe that notice-and-consent schemes are *also* inappropriate because, even though individuals might have the moral authority to consent, they cannot do so validly under current conditions, for instance due to lack of reasonable alternatives, a lack of information about the consequences of sharing information, and so on (Susser 2019).

We primarily mention these broader issues to set them safely aside, however. We shall not take a view on the legal adequacy of notice-and-consent schemes, and our attention shall strictly be on a moral argument that questions individuals' ability to give valid consent to data collection via pointing out third-party consequences of such transactions. Narrowly, our focus shall be on what follows from a commitment to the PDT.

3 Privacy dependencies and self-defense

In this section, we motivate our challenge to the PDT. We do so in two steps. First, we comment on the general relationship between a generic theory of permissible self-defense, on the one hand, and the PDT on the other. Next, we will discuss some cases offered by Véliz that is meant to illustrate the precise scope of the PDT. We then argue that it seems counterintuitive that even very mild forms of self-defense are permissible in at least some of these cases. This, we suggest, calls for a revision of the PDT.

Under normal circumstances, agents have non-defeasible duties to not act in ways that wrong or (wrongfully) harm others. As many have recognized, though, this picture changes when agents become subject to unjustified threats posed by others (for an overview, see Frowe and Parry 2022). As an example, imagine that A attempts to murder B. When faced with such a threat, B may permissibly engage in self-defense and harm—or even kill—the unjustified aggressor A, provided that doing so is a necessary and proportionate means to neutralize the unjustified threat. This idea is a perfectly general one, and it is widely endorsed (Lazar 2012). For our purposes, though, our aim is to highlight that the PDT picks out grounds for permissible defensive harm, provided certain further conditions are satisfied. Let's fill in the details. According to Seth Lazar (and many others), there are four basic criteria that must be satisfied for defensive harm to be permissibly imposed.

3.1 The standard theory of permissible self-defense

a. The defender must face an unjustified threat.
b. There must be some grounds to prefer the defender's interests to those of his target.
c. The force used must be proportionate to the threat averted: the threat must be of sufficient magnitude to justify that much force.
d. The force used must be necessary to avert the threat (Lazar 2012: 3–4).[5]

Advocates of the PDT must say that the first condition is satisfied in cases where one shares information about oneself that threatens revealing information about others. This is so, because endorsing the PDT implies that the conduct of A sharing *p* with C is an activity that is unjustified and an activity that consists of posing a threat to B. Here is why A's conduct is unjustified according to PDT: the thesis holds that A has a duty not to share *p* with C, so in the absence of any countervailing reasons, sharing *p* with C would be an unjustified action for A to perform. If you have a duty to abstain from α-ing, it could not be justified to α-ing.

Proponents of the PDT are committed to the claim that people pose a threat to others when sharing information because of the arguments they invoke to justify it. Let's go over the two defenses of the PDT we identified in the former section to see this. Begin with defense (1). Advocates of the PDT argue that when A shares *p* with C and thereby makes it possible for C to infer *q* about B, then A facilitates a violation of B's *right* to privacy. This is why there is a threat from A. It is an ongoing debate in the literature if inferring personal information about other people can constitute privacy rights violations, but several prominent theorists have argued it can (Rumbold and Wilson 2018; Munch 2021, 2022b).[6] If this view is indeed correct, then it means that C violates B's right to privacy when C infers *q* from *p*. But if so, one could suggest (as proponents of PDT in fact do) that not only does C violate B's right to privacy, A facilitates the rights violation by sharing *p* with C, making it possible for C to infer *q*. If A in fact facilitates the violation of B's right to privacy by sharing *p* with C, then the PDT implies that A poses an unjustified threat to B by sharing *p* with C.

What about defense (2)? Even if one denies that C violates B's right to privacy by inferring *q* from *p,* and consequently denies that A violates B's right to privacy by sharing

[4] Mills (2022) questions whether the fact that a data sharing choice has an impact on third parties is sufficient to undermine the sharer's moral authority to authorize the choice.

[5] See also Frowe and Parry (2022).

[6] See, however Mainz (forthcoming), for an argument to the contrary.

p with C, there is still an explanation for why the PDT implies that B may engage in self-defense. The explanation is that personal information inferred from the personal information of others is often used in ways that seriously harm people, for example by imposing economic burdens on them. If C uses *q* in a way that harms B, this may explain why A poses an unjustified threat to B (irrespective of whether any privacy rights are violated). So, the twofold defense advocates give for the PDT, in conjunction with endorsement of the thesis itself, forces them to say that the first condition of permissible self-defense is satisfied.

Now, what about the remaining conditions (ii)–(iv)? It is easy to imagine how they can be satisfied as well. Let us start with condition (ii). Generally, there are two competing views in the literature on what it takes for condition (ii) to be satisfied. On either of these views, it seems that condition (ii) is satisfied in the types of scenarios we are concerned with here. On the first view, condition (ii) is satisfied when the attacker is liable for imposing the threat in question (McMahan 2005; Thomson 1991). If what A does is indeed unjustified because A threatens B's right to privacy and threatens to impose harm on B, then A is presumably liable. On the second view, it does not matter if the attacker is liable or not. Even if the attacker is not liable, the defender always has an agent-centered prerogative to prefer her own interests over the attacker (Frowe 2008; Quong 2009). On this view, condition (ii) is straightforwardly satisfied too, since it is justified in any situation in which someone imposes an unjustified threat on B. We shall not engage in a discussion about which view is correct, since condition (ii) is satisfied on both views.

Whether condition (iii) is satisfied depends on what B's defensive actions against A consist in. Of course, there are things B could do that would be disproportionate considering the type of threat A imposes on B. For example, it would hardly be justified to kill A to prevent her from sharing *p* with C. But perhaps there are other things that would be proportionate. Plausibly, it would be proportionate for B to hack A's email account, hack A's social media account, or something such, to make sure that A does not share *p* with C. It might even be justified to break into A's apartment and destroy her device before she shares *p* with C. Our argument does not depend on specifying exactly which actions B is permitted to perform to prevent A from sharing *p* with C. It does not matter exactly where we draw the line, and it might be hard to do so crisply since it seems possible to imagine cases where there could be reasonable disagreement about whether the amount of force employed is proportionate or not. What matters is that there are some actions that B may permissibly do to prevent A from sharing *p*.[7]

[7] For helpful discussions of condition (iii), see Hurka (2005); Frowe and Parry (2022).

Finally, whether condition (iv) is satisfied depends on what it would take for B to prevent A from sharing *p* with C. If A would abstain from sharing *p* if B just asks kindly, then surely hacking A's email or something such would not be necessary to stop A from sharing *p* with C. But if hacking A's email is the only action that would prevent A from sharing *p* with C, then this would indeed be necessary. Sometimes there may be less intrusive ways to prevent A from sharing *p*. And sometimes this does not even involve directly interfering in A's actions, but rather in undermining the privacy dependency between A and B, or something such. For example, it might sometimes be enough that B just deletes their *own* social media account, or perhaps that B does what they can to boycott or lobby against C, so that C is not able to exploit the dependencies between A and B. In many cases, it may be impermissible for B to choose the more intrusive ways of preventing A from sharing *p*. But given the size of the data economy, and given the power that some social media companies have, it will often not be enough that B just stays off social media herself, or that she boycotts or lobbies against those companies. Thus, in many cases, it will be necessary for B to engage in quite intrusive acts of self-defense to block the threat from A.

We have now seen that endorsing the PDT may commit one to claims about when self-defense is justified (provided one is not ready to reject a common theory of permissible self-defense). This is an interesting finding by itself, as it speaks to the question of what people may do in the data economy to protect their privacy. But more importantly, the relationship opens the way for (indirectly) testing the intuitive plausibility of the PDT by checking if it gives rise to unsavory implications with regard to when self-defense is permissible. Here is what we have in mind. Recall for instance Bank Loan, a case where the PDT would say, because of privacy dependencies, that sharing information is impermissible which in turn enables us to infer that there would be an unjustified threat to third parties. If so, we have seen that (proportionate and necessary) self-defense to prevent this would be justified. If it turns out that this implication seems counterintuitive, we may infer at least one of two things. Either we should infer that the specific measures do not satisfy the conditions for permissible self-defense, *or* we should infer that the PDT is too expansive in that it picks out acts that should not, suitably interpreted, count as *unjustified* threats (both could be true, though). In this way, we can use intuitions about when self-defense would be permissible in response to threats to one's privacy to check the scoping of the PDT.

To frame our challenge right, want to stress that we find the concerns that motivate the PDT sound. Think for instance of a case in which a third party would incur a serious risk of harm were someone to disclose their personal information (e.g., having one's same-sex sexual preferences

disclosed in a society where bigotry is widespread). It seems right to say here that defensive harm to prevent disclosure of said information would be permissible. That is to say, we are quite sympathetic to the harm-based grounding of the PDT. But recall cases such as Genetic Information, Political Affiliation and Bank Loan. It strikes us as counterintuitive to infer that even very modest forms of self-defense (e.g., hacking people's computers to prevent disclosure of information, or physically preventing them from sharing their information) would be permissible in all cases that falls under these descriptions. Moreover, we do not think this is due to the suggested measures being disproportionate—surely, some force must be permissible if we are to believe that people sharing their information pose unjustified threats to others. Instead, we suspect that the better conclusion is that, unless we have heard more to the contrary, that these cases might well involve *threats* to third parties because of the existence of privacy dependencies, but not *unjustified* ones. To be even more specific, consider remarks by Véliz made on a case that structurally resembles Genetic Information:

> Let's imagine a friend (or maybe an enemy) gives you a home DNA kit as a 'gift'. Such kits are being sold for about £100. By mailing your saliva sample, you are giving away most or all of your rights to your genetic information. That means that companies like Ancestry can analyze, sell, and communicate your genetic information as they wish (...) it might be that you're willing to take these risks for yourself. (...) but what about your family? Your parents, siblings and children might not be happy to have their genetic privacy stripped away" (Véliz 2020: 77).

Véliz might be right that the family would not want this to happen, but we do not think that this observation alone licenses the desired conclusion or even an explanation taking the form suggested in the PDT, namely that we can infer that you should not share your information if that would lead to the disclosure of information that we would normally have privacy rights protecting against (e.g., disclosure of genetic information).

To appreciate this, notice that it might not be counterintuitive to say that it would be impermissible for the family to respond with mild forms of self-defense in this case. But we suspect this judgment might only be due to the case being underspecified with regard to what is at stake for the person intending to use the test kit. Specifically, one might think that such kits are merely used for a trivial benefit or "for fun". But let's instead imagine that the person uses kit to learn if they have a potentially lethal genetic disorder that, if present, warrants a risky treatment. In cases such as these, it would seem to us right to say that one would *not* be a fitting target of defensive harm. This seems intuitively right because the user of the test kit has a compelling justification for why they act in a way that ends up disclosing private information about others. In the presence of such a justification, defensive harm seems inappropriate because the threat is *justified*.

But that is hard to square with the PDT. Why? The PDT says that if there is potential disclosure of information that would otherwise violate people's right to privacy, then this disclosure would be impermissible. But if we believe that the mere fact that there was a threatened disclosure of information that would otherwise violate people's privacy rights would be sufficient to generate a duty not to share the information (as the PDT effectively holds), then its hard to see how this verdict can be avoided. It seems, we suggest, that we must *either* be prepared to reject the standard framework for when defensive harm is permissible (in order to avoid the implication that defensive harm is permitted in revised case where the kit is not just used for fun), *or* concede that the fact that the information being shared about third-parties of a kind they would normally have privacy rights against, does not fully determine the moral status of what the DNA kit-users ought to do. Or stated simpler, we might want to say that there is indeed a threat to the privacy of third parties, but deny that it is an unjustified threat.

What is the upshot of our challenge? If we are right, it cannot be true that whenever privacy dependency threatens to reveal a piece of information that would otherwise be protected by privacy rights, that one has a duty not to share one's information (as the original formulation of the PDT says). This is too strong. The best view here must be significantly weaker, given the pervasiveness of privacy dependencies in the data economy and given that it does not seem right to say that almost whenever somebody shares information about themselves, third parties are permitted in applying force to prevent them from doing so. Hence, we disagree with Véliz that it is possible to infer from "our interdependence in matters of privacy" to the conclusion that "no individual has the moral authority to sell their data." (Véliz 2020: 79). While it is hard to say what precisely should follow from the concerns that animate the PDT—a delicate matter we cannot settle here—we take ourselves to have shown that it might be less than has hitherto been suggested. The challenge, then, remains of saying precisely what follows from the fact that people's prospects for privacy are deeply interconnected in the data economy.

4 Four ways of resisting the implication

In this section, we discuss four objections to our argument. The strategy of each of them is to resist the implication of the PDT that it is sometimes permissible to harm other people to prevent them from sharing information about themselves. If you do not find the implication too problematic and find

yourself willing to bite the bullet and accept the implication if need be, then you should not worry if these objections are successful or not. If, on the other hand, you think the implication is a welcomed feature of the PDT and only adds to the plausibility of it, then you should hope that the objections are all unsuccessful. However, we suspect that many will find the implication unacceptable, and if you are one of them, you should hope that the objections are successful. Although we argue that all four objections are *unsuccessful*, we nevertheless think they are instructive to spell out. We begin with what we call the 'Liability Objection'.

4.1 The liability objection

The first objection proceeds from questioning whether people sharing information about themselves could truly count as liable targets for self-defense. Let's therefore call this objection the 'Liability Objection'. To motivate this objection, consider first that many believe, as we briefly touched upon above, that when we consent to sharing our own information in the data economy, our consent is mostly defective.[8] One reason for this is, as many have argued, that people are generally too poorly informed about the consequences of their data sharing choices (Susser 2019). A second reason that is mentioned is that people lack a reasonable alternative to sharing their information in the data economy (Susser 2019). A third reason stems from the claim that people are subjected to cognitive biases that invalidate their consent to sharing information (Solove 2012).

The reason why such claims might be of significance in the present context is that they suggest that when people share information about themselves that end up revealing information about others, they are not *culpably* posing an unjustified threat to others. Their inability to consent to data sharing practices, as the arguments mentioned above suggests, indicates that people are *blameless* in causing an unjustified threat. And if they are without blame when sharing data, this may in turn affect whether self-defense is justified.

We have two responses to this objection. First, people are not plausibly *always* blamelessly causing an unjustified threat to others in the relevant cases. To see this, recall the example discussed by Véliz (2020):

> Let's imagine a friend (or maybe an enemy) gives you a home DNA kit as a 'gift'. Such kits are being sold for about £100. By mailing your saliva sample, you are giving away most or all of your rights to your genetic information. That means that companies like Ancestry can analyze, sell, and communicate your genetic information as they wish (...) it might be that you're willing to take these risks for yourself. (...) but what about your family? Your parents, siblings and children might not be happy to have their genetic privacy stripped away" (Véliz 2020: 77).

Of course, *some* might fail to realize what they are really doing when they take DNA test kits. But saying that people cannot *in general* be sufficiently informed or uncoerced to consent validly to such arrangements, seems much too strong. It even seems too strong to say that people could not possibly be aware of the threat they are posing to others. Thus, we should not think that the person using the kit is necessarily causing a blameless threat to the effect that relatives could not be justified in imposing harm on the test kit user, i.e., by stealing their test kit or stealing the test sample in the mail. Moreover, even if we were to concede that such sharings of information *always* amounted to blamelessly causing a threat, then this would not show that self-defense is always ruled out. As many recognize, self-defense may be permissible even when targeted against people that are without blame causing an unjustified threat. A common argument for this view is that the *fairness* of the distribution of costs matters, and that it seems more fair to divert costs to those causing a threat, even if they do so *non*-culpably, rather than towards innocent victims that do not even *cause* a threat in the first place.[9] Under such a view, though somewhat controversial, we cannot block the inference that the PDT motivates self-defense by pointing out that those sharing information about themselves—and thereby impose costs on third-parties due to the existence of privacy dependencies—mostly do so without blame.

4.2 The causal contribution objection

Here is a second possible objection to our claim. It stems from reflection upon the (often small) causal contributions that information sharers are making to an unjustified threat *qua* their information sharing. Call therefore this the 'Causal Contribution Objection'. It seems right to say that in some cases in which people share information about themselves that could be used to make revealing inferences about others, their contribution is insignificant when each is considered in isolation. To appreciate this, contrast Identical Twins with Bank Loan. In Identical Twins, it is very plausible that the contribution Smith makes to the unjustified threat to Jones' privacy is causally indispensable. This is because Smith's activity is sufficient to impose a threat on Jones' privacy, and it is not easy to imagine that were Smith to abstain from sharing his genetic information, another person would cause

[8] See for instance Nissenbaum and Barocas (2016); Solove (2012); Mills (2022).

[9] Cp. Nozick (1974); McMahan (2005) for influential defenses of this idea.

a similar threat instead (let's imagine that they have no further family that could disclose the genetic information, to render the point vividly). By contrast, in the case of Bank Loan, it is plausible to imagine that were Peter *not* to give his information to the bank, some, perhaps many, other people would. In this sense, Identical Twins and Bank Loan differ because the threat is overdetermined in the latter case and not in the former. Using technical language, Bank Loan seems to share the structure of cases of 'collective harm'; cases where each single contribution that individuals make is negligible, but together they form a harmful pattern via the aggregation of information and privacy dependencies (Nefsky 2019). This at least seems to matter morally because it is less clear how it could be justified to impose harm in self-defense on people that make a, taken in isolation, trivial contribution to an unjustified threat.

There are at least two ways of interpreting this objection. On the first interpretation, the worry is that trivially contributing to an unjustified threat is *insufficient* to render one liable to bear costs to avert the unjustified threat. If successful, this objection would at best show that self-defense is not permissible in Bank Loan, not that it is not permissible in Identical Twins. Moreover, as many recognize, a slight contribution to an unjustified threat can be sufficient to be liable. To see this, consider:

> "*Bathtub*: Bad Guy wants to drown Victim in the bathtub and offers a $20 reward for helpers. Bad Guy holds Victim down, while 110 Helpers each pour one liter of water into the bath. 100 liters are sufficient to kill Victim." (Frowe and Parry 2022).[10]

It seems intuitively right that it would be permissible for Victim to kill at least one helper if this is necessary to escape. But since each helper's contribution is causally dispensable, it cannot be true that small or even trivial contributions cannot render one liable to harm for the sake of self-defense. By extension, there is nothing here preventing us from thinking that people sharing information about themselves could not become liable to be harmed in self-defense as a way of preventing the information sharing—as in Bank Loan.

The second interpretation of the objection points to the fact that in cases such as Bank Loan, merely preventing one from sharing their information will not avert the threat, since somebody else will plausibly contribute their information to make sure the inference can be made anyway. If so, it will not be permissible (because it will not be a necessary means to avert the threat) to prevent one person, with force, from sharing their information. This worry can be answered too, though. Our first response is that it only shows that one might be permitted to impose costs on many people, or, more precisely, enough people such that the threat is expectably averted. But conceding this does not undercut our claim the PDT implies that others may prevent you from sharing information about yourself in self-defense. Our second response is to say that in a complex system, such as the data economy, we cannot expect any individual to fully defend themselves against unjustified threats as a necessary condition for permissible resistance. However, this is consistent with saying that people are generally permitted in *fighting back*, even if their efforts are not guaranteed to be efficacious. Analogously, many would say that resisting an unjust regime may be morally permissible, even if one's resistance (that will likely impose harm on others) will not necessarily be successful. In fact, such an idea is far from alien to privacy theorists. Nissenbaum and Barocas (2014) argue, for instance, that it is permissible to impose harm on data collectors as a form of self-defense when their surveillance activities violate one's right to privacy. By similar lights, why not think that one is permitted to engage in self-defense against those that contribute to threats via privacy dependencies? Thus, the Causal Contribution Objection seems answerable by drawing on familiar resources from the literature on permissible self-defense.

4.3 The wrong target objection

Here is a third objection. Consider the fact that in all cases we have discussed, there is a data processor that is making the privacy-infringing inferences that lead to the activities that produce the downstream harms for data sharing. One might therefore think that it would always be more appropriate, perhaps out of concerns for fairness, to target the data processors instead of those contributing their information that facilitate wrongdoing. For example, it might be more appropriate to target Facebook instead of targeting specific Facebook users. Call this objection the 'Wrong Target Objection'.

We largely agree with the spirit of this objection. It may indeed be more appropriate to target the data processor. But conceding that we should sometimes give priority to targeting data processors rather than data sharers when we can, is fully consistent with saying that we could be justified in targeting both. And, unfortunately, it might turn out that in some cases it will prove ineffective to prevent the threat by targeting the data processor instead of the data sharer. To appreciate this point, consider the power asymmetries between individual users and companies such as Facebook. If one is concerned with Facebook making privacy violating inferences from the information other people have contributed, it might, given power asymmetries, be impossible to prevent Facebook from doing so. However, it might occasionally be more efficient to prevent people from contributing the data that enable Facebook's problematic activities

[10] See also Frowe (2014).

(for instance, in the case of Identical Twins). But, in any case, we can afford to admit that there might be a case for both targeting data processors and those that contribute the information that enable the wrongful inferences.

4.4 The objection from enforcement-inaptness

Here is a fourth and final objection. Call it the Objection from Enforcement-Inaptness. This objection takes as its starting point the claim that not all moral duties are what Barry and McTernan (2021) call 'enforcement-apt'. That is, some moral duties are simply not apt for enforcement. For instance, many reject that a spouse may forcibly prevent their partner from engaging in infidelity even though infidelity is morally wrong (given the existence of their relationship and) given that there are ways of preventing infidelity that would satisfy conditions (ii)–(iv) of the standard theory of permissible self-defense. It is debatable whether the alleged existence of non-enforceable duties poses a problem for the standard theory of self-defense, or if it merely shows that failing to comply with a non-enforceable duty implies that condition (i) is not satisfied. Either way, the alleged existence of non-enforceable duties reveals a possible objection: arguing that the duties that the PDT picks out are enforcement-*inapt.* If A's duty not to share *p* with C is enforcement-inapt, then the PDT can be true without implying that B is justified in harming A to prevent A from sharing *p* with C.

Here is our response to the objection. The first thing to note is that the kinds of considerations used in the twofold defense of the PDT seem like considerations that would typically allow for apt enforcement. That should be clearest in the cases where harm is threatened. To see that harm is at least sometimes a basis for enforcement-aptness in the context of privacy, consider a case where somebody threatens to disclose your same-sex sexual preferences in a society where homosexuality is sanctioned heavily both formally and informally. This strikes us a case where you would clearly be permitted to respond in self-defense to avoid this disclosure.

But such cases do not really cut to the core of our argument, since we did not stipulate harm being a central concern in the case with the DNA test kit we considered at length in the previous section. So, one might now reasonably want to ask: could it be that duties to not disclose other people's genetic information are only enforcement-apt provided a further harm would ensue because of the disclosure? That would be a problem for our argument, since this could provide an alternative explanation of why responding with self-defense does not seem appropriate in the DNA test-kit case where the kit is used to learn potentially important genetic facts about oneself.

However, upon closer inspection, we do not think this charge sticks. To see this, consider other cases of 'harmless' rights violations, such as stepping over somebody's lawn or making (harmlessly) use of somebody's property without consent. In such cases, even though we have stipulated the absence of harm, there clearly seems to be a basis for acting in self-defense (at least by imposing mild forms of harm), provided these are necessary. But if so, it is less clear why one should not be permitted to do something similar in the context of privacy. In fact, there also seems to be cases where 'harmless' threats of privacy violations intuitively merit self-defense. Suppose that you are naked in the shower, and the only way to prevent a stalker from observing you is by throwing a stone at him that will impose a moderate harm on him and scare him away. Such a move strikes us as intuitively permissible; but if that's right, then it's not clear that even harmless privacy violations would be non-enforceable.

Perhaps there is more to that needs to be said here. The question of the enforceability of privacy rights is a fascinating issue that have not been given much attention in the literature. But we cannot dive deeper into this matter here due to space constraints. So let us instead point out that some in the literature have in fact argued that privacy rights are enforceable (Nissenbaum and Brunton 2016; Munch 2022a). Hence, it would probably come as a surprise if harmless privacy rights violations were never enforceable and together with the examples we provided above, this strikes us a enough evidence to tentatively fend off the challenge from enforceability.

5 Concluding remarks

In this paper, we have presented a challenge for proponents of the PDT. Accepting the PDT, in conjunction with accepting the standard theory of permissible self-defense, implies that it is sometimes justified to harm people to prevent them from sharing information about themselves with others. While this implication adds important nuance to our understanding of the moral significance of privacy dependencies, we have also argued that it puts pressure on at least one version of how much we ought to do to restrict our data sharing in light of such dependencies.

Acknowledgements We thank Jørn Sønderholm for comments on a previous version of the manuscript.

Funding Open access funding provided by Royal Danish Library, Aarhus University Library. Work on this article has been generously supported by a Carlsberg Foundation Young Researcher Fellowship Grant awarded to Jens Christian Bjerring (Grant Number CF20-0257).

Data availability No data were collected or generated for this study.

Declarations

Conflict of interest There are no competing interests to be declared.

Ethical approval Not applicable.

Consent to Participate Not applicable.

Consent to Publish Not applicable.

References

Barocas S, Levy K (2019) Privacy dependencies. Wash Law Rev 95:555

Barry C, McTernan E (2021) A puzzle of enforceability: why do moral duties differ in their enforceability? J Moral Philos 19(3):229–253

Cohen JE (2012) Configuring the networked self: law, code, and the play of everyday practice. Yale University Press, New Haven

Fairfield J, Engel C (2015) Privacy as a public good. Duke Law J 65:385–457

Frowe H (2008) Threats, bystanders and obstructors. Proc Aristot Soc 108:365–372

Frowe H (2014) Defensive killing. Oxford University Press, Oxford

Frowe H, Parry J (2019) Wrongful observation. Philos Public Aff 47(1):104–137

Frowe H, Parry J (2022) Self-defense. In: Zalta EN (ed) The Stanford encyclopedia of philosophy. https://plato.stanford.edu/archives/sum2022/entries/self-defense/

Hurka T (2005) Proportionality in the morality of war. Philos Public Aff 33(1):34–66

Lazar S (2012) Necessity in self-defense and war. Philos Public Aff 40:3–44

MacCarthy M (2011) New directions in privacy: disclosure, unfairness and externalities. 6 I/s J Law Policy Inf Soc 6:425

Mainz J (forthcoming) Inferences and the right to privacy. J Value Inquiry

McMahan J (2005) The basis of moral liability to defensive killing. Philos Issues 15(1):386–405

Mills K (2022) Consent and the right to privacy. J Appl Philos 39(4):721–735

Mühlhoff R (2021) Predictive privacy: towards an applied ethics of data analytics. Ethics Inf Technol 23:675–690

Munch L (2021) Privacy rights and 'naked' statistical evidence. Philos Stud 178:3777–3795

Munch LA (2022a) Digital self-defence: why you ought to preserve your privacy for the sake of wrongdoers. Ethic Theory Moral Prac 25:233–248

Munch LA (2022b) How privacy rights engender direct doxastic duties. J Value Inquiry 56:547–562

Nefsky J (2019) Collective harm and the inefficacy problem. Philos Compass 14(4):e12587

Nissenbaum H (2010) Privacy in context: technology, policy, and the integrity of social life. Stanford Law Books, Stanford

Nissenbaum H, Barocas S (2014) Big data's end run around anonymity and consent. In: Lane J, Stodden V, Bender S, Nissenbaum H (eds) Privacy, big data, and the public good: frameworks for engagement. Cambridge University Press, Cambridge, pp 44–75

Nissenbaum H, Brunton F (2016) Obfuscation. A user's guide for privacy and protest. MIT Press, Cambridge

Nozick R (1974) Anarchy, state, and utopia. Basic Books, New York

Quong J (2009) Killing in self-defense. Ethics 119:507–537

Reidenberg J et al (2014) Privacy harms and the effectiveness of the notice and choice framework. In: TPRC Conference Paper, Fordham Law Legal Studies Research Paper No. 2418247

Rumbold B, Wilson J (2018) Privacy rights and public information. J Polit Philos 27(1):3–25

Solove D (2012) Privacy self-management and the consent dilemma. Harvard Law Rev 126:1880

Susser D (2019) Notice after notice-and-consent: why privacy disclosures are valuable even if consent frameworks aren't. J Inf Policy 9:37–62

Thomson J (1991) Self-defense. Philos Public Aff 20:283–310

Véliz C (2020) Privacy is power. Bantam Press, London

Zarsky T (2004) Desperately seeking solutions: using implementation-based solutions for the troubles of information privacy in the age of data mining and the internet society. Me Law Rev 56:13