



A Value-centered Exploration of Data Privacy and Personalized Privacy Assistants

Sarah E. Carter^{1,2,3} 

Received: 2 November 2021 / Accepted: 2 November 2022
© The Author(s) 2022

Abstract

In the current post-GDPR landscape, privacy notices have become ever more prevalent on our phones and online. However, these notices are not well suited to their purpose of helping users make informed decisions. I suggest that instead of utilizing notice to elicit informed consent, we could repurpose privacy notices to create the space for more meaningful, value-centered user decisions. Value-centered privacy decisions, or those that accurately reflect who we are and what we value, encapsulate the intuitive role of personal values in data privacy decisions. To explore how we could design for such decisions, I utilize Suzy Killmister’s Four-Dimensional Theory of Autonomy (4DT) to operationalize value-centered privacy decisions. I then utilize 4DT to help design a system—called a value-centered privacy assistant (VcPA)—that could help create the space for value-centered data privacy decisions using privacy notices. Using this 4DT lens, I further assess the degree that an existing technology, personalized privacy assistants (PPAs), use notices in a manner that allows for value-centered decision-making. I lastly utilize insights from the PPA assessment to inform the design of a VcPA, concluding that a VcPA could utilize notices to assist users in value-centered app selection and in other data privacy decisions.

Keywords Data privacy · Privacy notice · Privacy assistant · Autonomy · Values

✉ Sarah E. Carter
s.carter6@nuigalway.ie

¹ Science Foundation Ireland Center for Research Training in Digitally-Enhanced Reality (D-Real), Galway, Ireland

² Discipline of Philosophy, School of History and Philosophy, College of Arts, Social Sciences, and Celtic Studies, University of Galway, Galway, Ireland

³ Data Science Institute, University of Galway, Galway, Ireland

1 Introduction

Over the last four years, those of us browsing the Web from the European Union have become accustomed to a familiar sight—privacy notices, or “privacy pop-ups,” asking us to agree to a website’s privacy policy. These notices are a result of the General Data Protection Regulation (GDPR), passed in 2016 and implemented in 2018. Under the GDPR, informed consent is one of six legal bases¹ by which personal data can be collected and processed. This legal basis resulted in a flurry of privacy notices to elicit consent (Degeling et al., 2019).

While disclosures of this kind existed before the GDPR, the sheer volume of them has brought discussions surrounding the value and effectiveness of privacy notices back to the forefront. Privacy notices themselves are grounded in a long history of consent in other fields, especially bioethics.² Like consent forms for medical procedures or clinical trials (Beauchamp, 2011), privacy notices ideally aim to elicit informed consent from an individual.

Asking for consent has normative roots in respecting autonomy. Foundational documents of bioethics, such as the Belmont Report and the Declaration of Helsinki (Ryan et al., 1979; World Medical Association, 2013), stress the importance of respecting autonomy as a means to checking exploitative and manipulative practices.³ James F. Childress, co-author of the Belmont Report, further purposes that respect for autonomy requires that we not only refrain from manipulating (a negative duty), but also that we disclose information and foster autonomous decisions (positive duties) (Childress, 1990). These duties have been operationalized into informed consent—that is, facilitating voluntary choice by providing information that is comprehensive and complete (Beauchamp, 2011).

For privacy notices to be effective mechanisms of informed consent, then, they must allow for voluntary choice based on comprehensive information. Discussions surrounding the “privacy paradox” phenomena—or the observation that the choices we make when faced with a privacy notice do not match what we say our privacy preferences are—suggests that neither of these criteria for informed consent are met (Spiekermann et al., 2001). There are two general arguments that have been used to explain the privacy paradox, one in line with privacy decision-making as a rational, preference-based endeavor and the other postulating that humans are “flawed” thinkers who make questionable privacy decisions (Solove, 2020). The first argues that our privacy behavior, and not our stated privacy preferences, *truly* reflects how much

¹ The six legal bases for processing data are: consent; contract; legal obligation; vital interests; public task; or legitimate interests. These are specified in Article 6 of the GDPR.

² There are special considerations to medical research that may not be applicable to data collection situations—e.g., issues of dependency and vulnerability in healthcare situations introduce added consent concerns. There are, however, overlaps between bioethics and data ethics—especially as Big Data, smart-phone apps, and artificial intelligence enter the medical field. For examples, see: Jongsma et al., 2018; Klugman et al., 2018; Lucivero & Jongsma, 2018.

³ The traditional catalysts for the Belmont Report and the Declaration of Helsinki were the Tuskegee syphilis trials, where black men were left with untreated syphilis and observed, and revelations surrounding Nazi experimentation on prisoners during WWII. See: Brandt, 1978; Carlson et al., 2004.

we value privacy against, say, the good or service being offered. This is because we have been adequately informed by the privacy notice or policy and are acting in a rationale manner. In contrast, the second argues that our data privacy choices when faced with a privacy notice are not an accurate reflection of privacy's value but rather the result of human bias and heuristics—cognitive “tricks” or shortcuts that help us make fast decisions (Thaler & Sunstein, 2008; Tversky & Kahneman, 1974). Cognitive and behavioral psychology and its offshoot, behavioral economics, overwhelmingly support the latter view—and, indeed, few would agree that we are purely rational *homo economicus*⁴ when it comes to the decisions we make when faced with a privacy notice.⁵ This suggests that not only are we not informed by privacy notices, but that we are susceptible to manipulation through exploitation of heuristic-based thinking. To explore these challenges in more depth, four examples of this “flawed” human thinking, followed by a brief look at the lack of comprehension problem, are explored below.⁶

To begin, there are *framing effects*. We are more likely to consent to data collection if it is framed in nonthreatening ways (O'Neill, 2002; Thaler & Sunstein, 2008). In addition, we can be encouraged to consent by utilizing the *inertia bias*—or, the human tendency to stay with the default condition (Thaler & Sunstein, 2008). For example, pre-selecting “Agree to All” in privacy notices increases consent to online data tracking (Utz et al., 2019).

We can also be coaxed into a false sense of security by utilizing the *representative heuristic* and/or the *conjunction fallacy* (Lewis, 2017; Thaler & Sunstein, 2008; Tversky & Kahneman, 1974). Privacy notices usually link to privacy policies—long, complicated documents packed with legal terminology (Jensen & Potts, 2004). We may believe something to be true because it matches our mental images of what it should look like—in this case, a long legal document resembles a “strict” privacy policy. This is the *representative heuristic*. Closely related to the representative heuristic is the *conjunction fallacy*, or the tendency to have our predictions misled when flooded with truthful facts (Lewis, 2017; Tversky & Kahneman, 1974). Privacy policies can be written to selectively reveal or flood us with truthful facts in a way that can mislead our predictions because they remain tactfully vague on other aspects.

Indeed, service providers that collect data can utilize these design strategies to coax us into making the choices they want us to make. These designs, called “dark patterns” or “deceptive design patterns,” are ubiquitously used on privacy notices and have noticeable, measurable effects on data privacy decision-making (Brignull,

⁴ Term from Hoofnagle and Urban (2014). This text explores the influence and pitfalls of American privacy scholar Alan Westin's segmentation model, where a rational consumer reads privacy policies and makes decisions based on her preferences. There are three types of consumers according to Westin: “privacy pragmatists,” “privacy fundamentalists,” and “privacy unconcerned,” with the pragmatist in the majority. According to Hoofnagle and Urban, however, these “pragmatists” are hardly pragmatic at all, often making decisions with little or no understanding of the privacy protections in place.

⁵ There is also discussion around whether privacy is something that can or *should* be tradable for a good or service. For example, see Allen (2013).

⁶ For a more comprehensive overview, see Solove (2020).

n.d.; Gray et al., 2018; Mathur et al., 2019; Utz et al., 2019). To describe the consent given on these privacy notices as voluntary would be a stretch.

Now on to comprehension. While it is difficult to measure and results vary, it has been suggested that as low as 0.24% of us read online privacy policies (Jensen & Potts, 2004). Practically speaking, no reasonable person *could* read all of them—one study estimated that it would take 244 h to read all the privacy policies we see in one year alone (McDonald & Cranor, 2008). Like this challenge of *privacy policy fatigue*, too many privacy notices can cause *privacy notice fatigue*. This results in many of us deciding to simply “click through” notices rather than reading them (Ben-Shahar & Schneider, 2010; Schaub et al., 2015). Disturbingly, a study of online social networking services estimated that this “click through” rate could be as high as 74% (Obar & Oeldorf-Hirsch, 2018). Eventually, this notice fatigue can cause us to become “apathetic users”—those who decide to consent every time to a service’s data collection practices because we “no longer care” about our data privacy.⁷ We are, in sum, overwhelmed instead of informed by privacy notices.

These revelations from behavioral and cognitive psychology pose significant threats to the value of privacy notices. To address these challenges, research has been conducted to redesign and improve privacy notices to be more effective. Nearly every aspect of notice design—from timing, channel, modality, and control—has been extensively explored (Schaub et al., 2015). Particular interest has been paid to utilizing our heuristics to our benefit—that is, to “nudge” us beneficially. Nudging involves altering someone’s environment in such a way that encourages one decision, without forbidding alternative courses of action (Thaler & Sunstein, 2008). In the case of privacy notices, these nudges can be, for example, in the direction of better privacy (preserving) choices, or to increase comprehension/retention of a privacy policy (Almuhimedi et al., 2015; Calo, 2012; Kelley et al., 2013; Waldman, 2018). These interventions have ranged from presenting privacy permissions in a clearer format (“Privacy Facts”) when downloading a new smartphone application (Kelley et al., 2013) to informing users how much their location has been used (Almuhimedi et al., 2015). These privacy-preserving nudges on privacy notices are sometimes called, to contrast them to dark patterns, “bright patterns” (Grassl et al., 2021).

By utilizing notices to encourage us to “preserve our privacy,” however, bright patterns can undermine the original goal of privacy notices—to respect autonomy and prevent manipulation. This is due to the nature of nudges themselves. There remains a fundamental, ongoing discussion concerning whether nudges are inherently paternalistic or manipulative, linked to larger discussions around autonomy. Of particular interest is the concept of volitional autonomy—or that our actions should reflect our authentic desires. Volitional autonomy can be traced to analytical philosophers Harry Frankfurt and Gerald Dworkin. According to Frankfurt’s original

⁷ There is a caveat here that is worth mentioning: The term “apathetic user” is meant to capture those who would prefer to be more data protective but feel overwhelmed by notices to the point of “no longer caring,” or apathy. Someone who reflects upon their data privacy choices and decides to give their data away “always” because they, say, believe that sharing their data will stimulate technological progress and improve their user experience are not meant to be captured as apathetic users. However, I would argue that few of us “click through” notices because we fully endorse what we are doing.

theory of free action, to act freely—autonomously—we must act in accordance with our second-order volition, or, in other words, we must desire what we desire (Frankfurt, 1971). Dworkin similarly argued that autonomy is our ability to reflect and endorse “first-order” desires in accordance with our “higher-order” values (Dworkin, 1988b). Because those using nudges encourage the actions they desire the person nudged to desire, and not what the person being nudged necessarily desires to desire, the person nudged may act in a manner that they do not endorse (Hausman & Welch, 2010; Schmidt & Engelen, 2020). Put a bit more clearly, bright patterns encourage us to preserve our privacy by making us want to preserve it, even if we don’t “truly” want to. This violates our autonomy.

This concern around nudges and user autonomy more generally has been raised in other fields. In the field of psychology, for example, a similar conception of autonomy to that of Frankfurt’s and Dworkin’s has been operationalized by psychologists Richard Ryan and Edward Deci into their self-determination theory (SDT). SDT defines autonomy as self-endorsement of an action according to one’s values (R. M. Ryan & Deci, 2004). Empirically, studies utilizing SDT’s Internal Locus of Causality, or I-PLOC, measurement suggest that even small nudges can undermine someone’s autonomy (Arvanitis et al., 2020). This has implications for bright patterns, whose use—however well-intentioned—could cause an I-PLOC shift, again violating, rather than respecting, autonomy.

We are stuck at an impasse when it comes to how to design and utilize privacy notices. Insights from behavioral psychology suggest that informed consent via notice is not plausible. Instead, we are at worst, uninformed, and at best, overwhelmed when encountering a notice. We are manipulated and framed to “consent” our data away. Some have aimed to use privacy notices to nudge us to make “better” privacy choices, but this has possible negative implications on our individual autonomy. Perhaps, as many scholars argue, we should not use privacy notices at all.

I don’t believe this is the right question to be asking. Instead, we could try to imagine a role for privacy notices outside that of eliciting informed consent. Perhaps it is not the notices themselves that are the problem, but a mismatch between the task and the tool. This argument has been made by Susser (2019), who suggests that privacy notices could perhaps promote individual decision-making by increasing awareness of the current data privacy landscape; helping identify potential privacy concerns; and assessing our legal rights, where applicable.⁸

In this paper, I would like to take this one step further, by marrying Susser’s intriguing idea to explore notice outside traditional notice-and-consent to notice’s original goal of respecting autonomy. While we must not prioritize respect for autonomy above all else and still balance it with other values relevant to data privacy—values such as accountability, transparency, and trust (O’Neill, 2020; Waldman,

⁸ See Susser et al. (2019, pg. 38): “If the problem with notice-and-consent as a whole is that it fails to facilitate and respect individual agency over data, then we ought not to deprive ourselves of even flawed and partial mechanisms for strengthening such agency.”

2015, 2018)⁹—respecting autonomy is still a worthwhile aim for data privacy decisions. Like in bioethics, respecting autonomy in data privacy decisions still provides a critical check to exploitation and manipulation despite informational power asymmetries—it is, at the very least, the first line of defense (Susser, 2019; Susser et al., 2019). There is also a sense of disempowerment and an increasing learned norm of simply “giving up” on protecting our data privacy. In the case of those who reside in liberal democracies, this disempowerment is fundamentally out of synch with the central tenants of our governments.¹⁰

To respect autonomy, then, we could try to utilize the insights of behavioral psychology and accounts of autonomy in philosophical literature to understand autonomy in a manner that doesn't rely upon informed consent. To be clear, I am not aiming to make regulatory recommendations or assess the GDPR, which, as the previous paragraphs demonstrate, likely will need to atone to the lessons of behavioral psychology just like notices. I leave this task to other scholars.¹¹ Instead, I am intrigued by Susser's call to find other uses of notices beyond traditional informed consent.

I am particularly interested in conceptualizing autonomy in a manner that better captures the intuitive role of personal values in data privacy decisions—what I call *value-centered* privacy decisions. This is because re-purposing notices for value-centered privacy decisions could retain the original normative underpinnings of notice—to respect autonomy—while also promoting more meaningful data privacy decisions and an overall better experience. Because of this emphasis on personal values, I will also not be exploring here the collective implications of privacy disclosure or the social value of privacy.¹² Similarly, value-centered privacy decisions are not meant to encapsulate discussions around broader, public values that may be relevant for governments when constructing policies with data privacy implications.¹³ I do not intend to dismiss these discussions, but rather to complement them with one re-focusing on the individual experience of data privacy choices—those small, often unsatisfying privacy decisions that we make daily as we interact with

⁹ Trust—and placing it well—is a pressing concern for data privacy decisions. For example, Onoara O'Neill (2020) notes several problems with trust and accountability in the digital age. Technology allows new intermediaries to control online content opaquely, and it is difficult to decide whether they are trustworthy. This results in unrealistic accountability mechanisms, like data privacy policies and excessive notices.

¹⁰ I particularly appreciate how this is stated in Susser et al., 2019, pg. 8: “Autonomy is in many ways the guiding normative principle of liberal democratic societies. It is because we think individuals can and should govern themselves that we value our capacity to collectively and democratically self-govern.”

¹¹ This has been done quite well elsewhere—in particular, Frederik Zuiderveen Borgesius devoted an entire PhD to this subject. While this was completed in pre-GDPR (EU Data Protection Directive) days, many of his analyses regarding informed consent remain relevant to the GDPR. See Zuiderveen Borgesius (2015).

¹² For an example of a discussion exploring the social value of privacy, see Roessler and Mokrosinska (2013).

¹³ Examples of these discussions can be found around COVID-19 tracking applications. The purported benefit to public health was considered in terms of efficacy and uptake of the apps (Luciano, 2020; Morley et al., 2020); justice, equity, and solidarity with vulnerable subpopulations (Hendl et al., 2020); and civil liberties and surveillance (Kitchin, 2020), to name a few.

digital technologies. In addition, once we understand autonomy for notice using value-centered privacy decisions as a starting point, we could utilize dynamic privacy technologies to translate this into practice. We could, in other words, instead of designing for more *privacy-protective* decisions, design for *value-centered* privacy decisions.

To this end, I will firstly define what is exactly meant by “value-centered privacy decisions.” To accomplish this, I will draw from interdisciplinary literature as well as invoke our own intuition. I will then operationalize value-centered privacy decisions using an autonomy lens—in particular, Suzy Killmister’s Four-Dimensional Theory of Autonomy (4DT) (2017).

In the second half of the paper, I move on to how we could use notices and dynamic privacy technologies to create a space for value-centered data privacy decisions. To do this, I will utilize 4DT to help design a system—called a value-centered privacy assistant (VcPA)—that could help create that space by using privacy notices. The VcPA is conceptually rooted in an existing technology, called personalized privacy assistants (PPAs), which are machine-learning systems that help users manage their data privacy (Liu et al., 2014, 2016; Warberg et al., 2019). I therefore firstly will utilize 4DT to assess to what extent value-centered decision-making is supported by existing PPA technology. I then will utilize insights from this analysis to inform the design of a VcPA to help users when they are deciding whether or not to download an app on the Apple or Google Play Store. I lastly conclude that a VcPA with these modifications could effectively utilized notice to assists users in value-centered app selection and in other data privacy decisions.

2 Notice Revisited: Promoting Autonomous, Value-centered Privacy Decisions

Privacy notices are not sufficient to promote informed privacy decisions. In this section, I explore how they could instead be utilized to create the space for us to make autonomous, value-centered privacy decisions—effectively designing for autonomy, rather than for privacy, in privacy decision-making.¹⁴ I firstly outline how values are involved in privacy decisions before linking this understanding to autonomy.

2.1 Individual Data Privacy Decisions (Ideally) Reflect User Values

Value-centered privacy decisions are a means of capturing the intuitive role of personal values in data privacy choices, with the aim of making the individual experience more meaningful. Value-centered privacy decisions can be understood as those

¹⁴ Exploring ways of promoting autonomy in technological design is not a new concept. It has, for example, been explored in the field of human–computer interaction (HCI) and is a central component of value-sensitive design (VSD) (see Peters et al., 2018; Friedman & Nissenbaum, 1997; Batya Friedman, 1996). My approach, however, is slightly different—it links personal values to data privacy decisions, which in turn, are further conceptualized and operationalized using autonomy. This is explored more in Sect. 2.

that result in ends that accurately reflect our personal values. Consider your home—the furniture you have there; the art you put on the wall; and who you let in should ideally be a reflection of who you are and what you value. We can imagine, however, instances when this won't be the case—perhaps your local furniture store doesn't carry the color of furniture you like and other stores are too far away for you to reasonably travel to. Similarly, your smartphone is like your “digital home”—what apps you download and who you allow to access your data should ideally reflect your values.¹⁵ However, hurdles such as dark patterns may restrict your privacy choice to the point that it may not reflect your values. Individual privacy decisions can therefore be evaluated to the extent to which they result in ends that accurately reflect the individual's personal values.

Many disciplines have explored this relationship between privacy and values more broadly, albeit in different contexts and with different emphases, and I have drawn from this multidisciplinary literature to inform this conceptual link of individual data privacy decisions and personal values. On the legal side, Daniel Solove (2002, 2007) has conceptualized privacy (more generally) as many related items that can be encompassed under a common heading without necessarily having a single theoretical basis. He further puts forth that privacy is instrumental in that it allows an agent to protect or promote valuable ends. Seeing individual privacy decisions as value reflection retains his idea of both value plurality and instrumentality. Here, however, value reflection emphasizes less the ends that are brought about than which personal values are ultimately reflected by the user achieving those ends.¹⁶

Values as ends in themselves is drawn partly from value-sensitive design (VSD), which proposes that technology embeds and expresses values (Friedman et al., 2008). I say in part because an individual's data privacy decision has consequences that could be seen, ideally, as a reflection of her *personally-held* values, rather than greater stakeholder values that concern a designer utilizing a VSD approach; it is, one could say, VSD with an emphasis on the individual experience. Value-centered privacy decision-making also incorporates an aspect of *information flow*, or, in this case, the data that is being collected as the result of user's privacy decision. This a central idea of philosopher Helen Nissenbaum's contextual theory of privacy (2004), which puts forth that privacy can be understood as *appropriate information flow* that match the context-specific norms. Here, again, value reflection is not interested in generally consensual norms but rather the individual's specific values and the degree to which they are accurately reflected following a data privacy decision.

¹⁵ Privacy “the value” is here understood as an instrumental value, valuable for other terminal values that “being private” brings about. Conceptualizing privacy in this manner is helpful for encapsulating the myriad of personal values—some compatible with privacy and some not—that are relevant when making a privacy decision. In other words, we can consider not only values associated with *privacy-preserving* behavior (such as personal security), but also other values that may be associated with *privacy-violating* behavior (such as the social value of sharing photos on Facebook).

¹⁶ Ibid.

2.2 Ideal Value Reflection (and Designing for It) Requires Autonomy

As stated in the introduction, designing for value-centered choices can be linked to the original goal of notice—respect for autonomy. This is because value reflection cannot occur if autonomy is frustrated. In order for value-reflection to be “ideal”—that is, an accurate reflection of someone’s values—a privacy decision must be sufficiently unhindered by external forces such as dark patterns. She must also be sufficiently engaged with the privacy decision to make a conscious, value-centered decision. We are, essentially, designing for autonomy—with the goal of providing more meaningful privacy decisions that reflect who someone is and what she values.

Furthermore, if we wish to explore ways of supporting value-centered data privacy decisions, we require a means of assessing to what extent one’s data privacy choices meets this “ideal.” We need, in sum, to define and identify conditions for autonomy in data privacy decisions.

To do this, we can look to the wealth of autonomy literature in philosophy to identify a concept autonomy that fits our context—individual data privacy decisions. To fit our purposes, the conception must meet three criteria.

Firstly, it must have *personal values* at its core to be harmonizable with the idea of value reflection.

Secondly, it must be *reasonably systematic* and practical in order to assess data privacy decisions.

Thirdly, because we are looking at individual data privacy decisions, it must be a theory of *personal autonomy* that is still able to account for the apathetic user phenomena (described in the introduction). We are not aiming to assess how the decision whether or not to disclose our data (our data privacy decision) affects our overall autonomy or greater democratic processes, but rather, how autonomous the decision to disclose our data is.¹⁷ In addition, we are not investigating whether what someone values and acts on in her privacy decision is the most morally justifiable one, but whether it is a reflection of her values. We are, therefore, not looking for a conception of moral or Kantian autonomy, but one of personal autonomy.¹⁸ At the same time, there must be some normative incentive to *not* become an apathetic user by succumbing to notice fatigue—otherwise, there will be no opportunity to make a value-centered choice. This final requirement—with all its nuances—is the trickiest of the three to meet.

Suzy Killmister’s Four-Dimensional Theory of Self-Governance (abbreviated as “4DT”) meets these criteria. 4DT divides autonomy into four dimensions—*self-definition*, *self-realization*, *self-unification*, and *self-constitution* (discussed in greater

¹⁷ For examples of these discussion, see Cohen (2013), who discusses the role of privacy in setting boundaries from external influences in order to define who we are and what we want, or Zuboff (2019), who explores the danger data disclosure and behavioral modeling pose to our agency and the greater democratic process.

¹⁸ This is not to say that privacy decisions cannot have moral weight, but merely that it is not what we are investigating here. For an example of moral duties towards privacy disclosure, see Allen (2013), where she draws on Kant to argue that we have a duty toward ourselves to protect our privacy out of respect for ourselves.

detail later in this section) (Killmister, 2017). To the first criteria, self-definition has links to values—it involves forming commitments on how to be and act in the world which, in turn, are clustered into values. 4DT also meets the second criteria: It categorizes autonomy into four distinct, accessible, and practical dimensions.

4DT can also account for the final, and most difficult, requirement—we require a theory of personal autonomy that also discourages apathetic choice. To the first criteria, 4DT is indeed a theory of personal autonomy—it is not concerned with whether the ends that are brought about are necessarily the morally correct ones. To the second criteria, it can also capture the “apathetic user” phenomenon. Normativity for 4DT is derived from an agents’ ability to take on or reflect upon their own commitments—that is, we must be committed to something we should do or become (Killmister, 2017). These agency requirements are captured in *self-constitution*. Thus, someone who refuses to take on any commitments or intentions whatsoever—such as an apathetic user—cannot be said to be highly self-constituting. Fitting this last part can be a challenge for closely related theories of personal autonomy and is, ultimately, where 4DT stands out. Returning to Dworkin, for example, if the “apathetic user” decides to take on no commitments concerning privacy decisions after the highest order of reflection, it could be considered autonomous (Dworkin, 1988a).¹⁹

Having established that 4DT meets the criteria for exploring value-centered data privacy decisions, we can now examine it in more depth; operationalize its dimensions in the context of data privacy decisions; and ultimately, in the next section, utilize it to design for value-centered privacy decisions.

As mentioned above, in her book, *Taking the Measure of Autonomy: A Four-Dimensional Theory of Self-Governance* (2017), Killmister maps the complex landscape of autonomy into four different dimensions. The first dimension, *self-definition*, is concerned with personal identity: Self-definition assesses the level of internal consistency between the goals, beliefs and values that make up our personal identity. While self-definition is not particularly relevant to the *act* of privacy decision-making, it does help us better visualize how values are involved and reflected in the privacy decision-making process. As outlined in Fig. 1, our self-defining attitudes—beliefs, goals, and values—commit us to act a certain way.²⁰ For example, perhaps I believe that too much screen time is bad for my health. I am therefore committed to spend less time on the computer. Similar commitments can then be clustered together into values. Perhaps I also have the goal of running a marathon. This commits me to run every day, which could be clustered with less computer time to say I value my health.

The second dimension, *self-realization*, is concerned with practical agency: Self-realization assesses the level to which our conclusion from practically deliberating

¹⁹ There are, however, instances where someone may act in a manner *similar* to an apathetic user while still being self-constitutive. For an example, please see footnote 7.

²⁰ Self-governing policies can also influence self-defining attitudes and have downstream effects on out privacy preferences. Self-governing policies dictate what and how to “believe, plan, and value” (Killmister, 2017, pg. 22).

aligns with our intentions (*internal self-realization*), and the degree to which our actions align with our intentions (*external self-realization*). Returning to our example, let's say I am considering downloading a running smartphone app that would like to access my health data. I deliberate and determine based on my values and, because I value my health, conclude that I best ought to allow it access. I form the intention, or in this case, a *privacy preference*, to allow the app access to my health data. I then act on this and allow it access—that is, make a *privacy decision*. This upholds self-realization, as I deliberated, formed an intention, and acted coherently.

The final two dimensions are concerned with the relationship between personal identity and practical agency. *Self-unification*, whether our actions are consistent with their personal identity, and *self-constitution*, our ability to take on and form commitments. In the case of downloading the health app, self-unification is upheld because my action to download the running app and give it access to my health data is consistent with valuing my health. Self-constitution is also upheld because I am able and willing to form conclusions and intentions concerning whether or not to download the app.

Applying 4DT to this context—individual data privacy decisions—comes with a few notable caveats. These concern self-unification and self-constitution. Firstly, self-constitution originally pertained to someone's *overall* willingness to take on commitments. She could, then, still be considered highly self-constituting if her willingness is present in a *range* of areas, even if this willingness does not extend to *all* circumstances or topics. In this case, I am utilizing 4D as a tool to analyze autonomy constrained to a specific circumstance—individual data privacy decisions—and am treating self-constitution as *the ability to take on any or new commitments pertaining to data privacy decisions*. Secondly, self-constitution can also apply to both the domains of personal identity and practical agency. It encompasses the degree to which someone forms self-defining attitudes; the degree to which she can practically deliberate and form intentions; and the degree to which she is able to coordinate the two. When considering how self-constituting she is with respect to data privacy, self-defining attitude formation and its (continued) coordination with practical deliberation is a longer-term process and not one that is made at any one privacy decision point; thus, self-constitution for our purposes will be primarily concerned with the degree to which someone is able to practically deliberate and form intentions. This brings me to the third and final caveat: Self-unification is not only concerned with

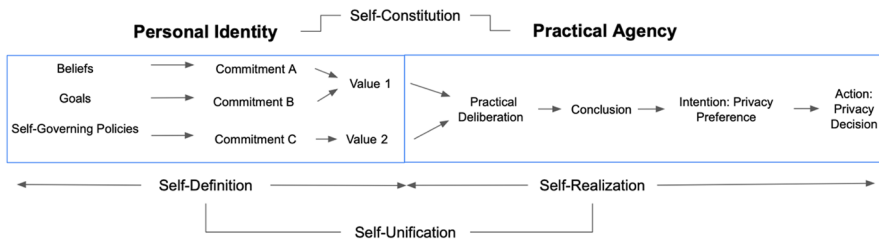


Fig. 1 Four-Dimensional Theory of Self-Governance (4DT) as it pertains to individual data privacy decision-making

ensuring that one's actions match their personal identity, but also the extent to which the conclusions/commitments generated during practical deliberation can inform transformation (a change in personal identity). Again, this is tied into self-defining attitude formation and is not something that is made at one point in time. Because of this, an analysis of data privacy-decision making in terms of the self-unification dimension will focus on the overall coherence of someone's actions to her personal identity.

3 Personalized Privacy Assistant Technology for Value-centered Privacy Decisions

So far, we have defined what value-centered privacy choices are and operationalized them using autonomy, and now, we can start designing for value-centered decisions. In this section, we will utilize the dimensions of 4DT to systematically assess the degree to which current personalized privacy assistants (PPAs) use notices to create space for privacy decisions that reflect our values (summarized in Table 1). In addition, we will also be mindful of insights from behavioral psychology that have undermined the value of privacy notice in the past, harmonizing them with 4DT where applicable. This analysis of PPAs is done to inform the design of a value-centered privacy assistant (VcPA) for smartphone app selection. Privacy assistants use dynamic and selective notices to improve the user privacy decisions, providing us with an excellent starting point for exploring notice use to create value-centered privacy decisions and eventually conceptualize a VcPA.

Table 1 Current personalized privacy assistant (PPA) evaluation using the Killmister's Four-Dimensional Theory of Autonomy (4DT) and suggested modifications for a value-centered privacy assistant (VcPA) for smartphone app selection

Challenge	Relevant Dimensions of Autonomy	Do Current PPAs Address this Issue?	Suggested Modifications for a Value Centered, Smartphone Privacy Assistant (VcPA)
Notice Fatigue	Self-realization Self-unification Self-constitution	In part: Personalized and selective notifications help prevent <i>notice fatigue</i> and promote more value-consistent privacy choices through self-binding, but do not best prevent double-binds	Keep selective notices, but base profiles on user values rather than privacy preferences alone to help minimize double-binds
Inertia Bias	Self-realization Self-unification	No: the <i>inertia bias</i> makes it difficult for users to change from their initial privacy settings	An <i>exploratory process</i> , perhaps using exploratory notices, that does not encourage one download choice over another
Lack of Controls	Self-realization Self-unification	No: because of the <i>limited granularity problem</i> , a user's ability to realize their commitments and act on them is hindered	<i>Suggesting alternatives</i> to quickly link users to similar applications whose data collection practices better align with their values

3.1 A Note on Context

I will firstly assess smartphone PPAs as they currently exist—that is, assisting users with privacy settings *after* downloading the application. I will then utilize this information to inform the design of a VcPA to assist a user with one specific privacy choice—when selecting a smartphone application. While there are many different individual privacy decisions that could be explored and could benefit from value-centered autonomy-enhancing privacy notices, it is not feasible to explore them all. The focus of this paper is when users are deciding to download smartphone applications on the Apple’s App or Google Play store, and there are a few reasons why this particular privacy decision point is of interest.

While it has been argued that a per-app permission basis is not a helpful context to explore for privacy interventions,²¹ there are several reasons why looking at the decision whether to download an application is appealing and could complement existing “post-download” interventions. To start, there is something to be said about empowering users to manage data collection in line with their values by preventing their data from being collected in the first place. While agreeing that the centralization of data collection into a select few third-party libraries cannot be combated with current permission choices, targeting the decision to download could take into account this third-party sharing issue by preventing value inconsistent data sharing to a library in the first place. It also provides an extra layer of control over a user’s personal data that could further complement—rather than contradict—other initiatives. Improving control in a manner that increases value reflection at the decision whether or not to download an application could also increase market pressure for improved products that better align with user values.²² In this manner, value-centered application selection could serve as the first line of defense against value-inconsistent app data collection practices and encourage more consistent applications to be made in the future.

3.2 Evaluating Personalized Privacy Assistants (PPAs)

We can now turn to personalized privacy assistants (PPAs): what they are, and why explore them as a means of facilitating value-centered smartphone app selection choices. Briefly, PPAs, currently under development by a team at Carnegie Mellon University,²³ would be machine learning assistants that personalize and automate

²¹ For example, Chitkara et al. (2017) have argued that looking to improve per-app permissions is unproductive because 70% of app data access requests are by the same 100 third party libraries. Instead, they created ProtectMyPrivacy—an application that provides users with the ability to control which libraries as well as apps have access to their data. In addition, cross-app tracking has also been seen as a major issue, leading to Apple to release the “App Tracking Transparency” feature, which requires smartphone apps to get permission from users to track their activity across many applications, with iOS 14.5 (Apple, 2021).

²² Susser (2019) makes a similar argument—privacy notices that disclose a company’s data practices, even if incomplete, could encourage them to meet social norms concerning privacy.

²³ Also see The Personalized Privacy Assistant Project (<https://privacyassistant.org>).

privacy choices for a user. The team has explored PPAs to help manage user privacy in the Internet of Things (Das et al., 2018) and smartphone applications (Liu et al., 2016), to name two. Personalized assistants could also utilize a variety of approaches to help a user manage their privacy, such as privacy-preserving nudges or semi-personalized setting recommendations based on user preference profiles (Liu et al., 2014; Story et al., 2020; Warberg et al., 2019). For example, Liu et al. (2016) designed a smartphone PPA where users got personalized recommendations for their privacy controls on their Android smartphone. They firstly developed privacy preference profiles based on a dataset of Android user privacy settings and then conducted a user study of the PPA. During this study, participants took a dynamic, short quiz on their privacy preferences to sort them into a privacy preference profile. Based on their profile, they were then given recommendations on how they could change their privacy controls on their phone.

Conceptually, the development of PPAs has been fueled by a desire to help users make the best privacy choices for them in the current digital privacy environment. In particular, PPAs seem to be focused on overcoming one aspect of this larger issue—user notice fatigue—something an intervention for smartphone application selection would also want to avoid (Liu et al., 2014). As Florian Schaub et al. have previously described (2015), determining the proper amount of notices is exceedingly difficult—too many notices causes users to simply “click through” them, and deploying them at too little a frequency doesn’t provide the user with adequate information to make an informed decision. Through dynamic, personalized recommendation, PPAs could overcome the notice fatigue concern by only presenting notices to the user that are relevant to them. This makes PPAs an attractive technology to help us make more autonomous, value-centered privacy decisions, where notice fatigue remains a challenge. By utilizing 4DT, we can explore PPAs to assess to what extent they are successful at dealing with the issue of notice fatigue by addressing the 4DT failures that constitute notice fatigue. We can also use 4DT to identify areas of improvement to inform the design of a value-centered privacy assistant (VcPA), a system to help use make app choices in a manner that best reflects our personal values.

3.2.1 PPAs and Notice Fatigue

Not surprisingly, from the standpoint of 4DT, PPAs are mostly successful at addressing notice fatigue by utilizing *selective* (rather than blanket) notifications (Liu et al., 2014, 2016; Warberg et al., 2019). The phenomena of notice fatigue can be understood as a failure to self-realize, self-unify, self-constitute, or a combination of these. We can consider failures to self-realization, where someone “just clicks through” *all* notifications despite an intention not to; self-unification, where this “click through” action does not match her personal identity and her values; and, in more extreme cases, self-constitution, where an “apathetic user” becomes so overwhelmed by privacy notices that she stops forming new intentions and commitments about her data privacy *at all*. In addition, challenges arise when two values come into conflict with

each other. For example, many of us, for a variety of reasons, value efficiency. We can, for example, imagine someone who values self-determination, control, and efficiency. In the current privacy notice landscape, her values will be at odds with each other—she will be in a *double bind*,²⁴ and no matter how she acts, she will act contrary to her values (fail to self-unify).

PPAs also act as a form of self-binding, that is, holding a user to follow through on her intentions. They do this by utilizing selective notifications to prompt her to re-consider her privacy decisions that are at odds with her preferences. This, in turn, may help her act more consistently with her personal identity. By selectively “slowing down” the user in a manner tailored to her individual privacy preferences (Kahneman, 2011), she has further opportunity to pause and re-consider an intention that is at odds with the conclusion of her practical deliberation.²⁵ This “slowing down” can be considered a beneficial form of friction, or disruption of a goal-focused mindset. While friction was originally considered something to be minimized in technological design, its positive aspects such as its link to slower, more mindful user thinking have received recent attention (Cox et al., 2016; Green, 2022).²⁶ In the field of data privacy, friction has been proposed as a means of encouraging reflection and attention to one’s own underlying beliefs and values when engaging in a data privacy decision (Terpstra et al., 2019). Friction in the case of PPAs can therefore be understood as a form of self-binding that helps the user self-realize and self-unify by triggering mindful reflection.

PPAs also introduce this friction in a manner that is more observant of what the user desires. I would argue that over-generous or designer-selective use of friction can cause problems. While these cases would encourage user reflection and, by extension, autonomous choice, the challenge here is that the *person designing the friction decides* when the user should slow down and be reflective, rather than when it would be most beneficial to the user to do so. This could either result in the user only making conscious decisions when the designer thinks that she should, or a well-intentioned designer may use friction too liberally as to cause notice fatigue. By selectively notifying users *based on what she thinks is best for her*, PPAs are theoretically better at assisting users without slipping into these issues.²⁷

Considering now the challenge of double binds, the PPA does help alleviate tensions between values—but only to a certain extent. In terms of a double bind concerning efficiency and other values described previously, she may receive an excessive number of selective notifications, and there may be instances where she receives

²⁴ I am using Killmister’s terminology here, from Part II: Applications of *A Four-Dimensional Theory of Self-Governance*, where she explores, among other aspects, the autonomy of the oppressed, and how they are often forced into double bind situations.

²⁵ Here, by “slowing down” I mean the shift from fast to slow thinking as described by Kahneman (2011). “Fast” System 1 thinking is automatic, mindless, and ripe with heuristics and biases. “Slow” System 2 thinking is conscious, deliberate, and mindful.

²⁶ Empirical studies have supported this idea, such as Mejtoft et al. (2019), where well-placed friction to increase mindfulness when interacting with a mobile application also increased user satisfaction.

²⁷ However, as *machine learning* assistants, PPAs need to address additional concerns (such as being sufficiently transparent). I discuss some of these issues in Sect. 3.4 when considering designing a PPA-like system for app selection, called a value-centered privacy assistant (or VcPA).

a notification that is not needed, and therefore, inefficient, given the large number of different possible privacy preferences and situations. Ideally, the notice's selectivity would be tuned to not overwhelm the user and lead to apathy, but the number of privacy preferences and situations could make this tuning exceedingly difficult. Thus, while PPAs can help with double binds in part, a better solution may be possible.

3.2.2 Other Challenges: Inertia Bias and Privacy Controls

Similar to the issue of notice fatigue is the inertia bias. As mentioned in the introduction, this cognitive bias makes it difficult for users to change from their initial privacy settings—even if they wish to (Thaler & Sunstein, 2008). In the case of PPAs, this manifests as a failure to update and change their privacy profile, even if they are given the ability to do so. Returning to the Liu et al. PPA paper (2016), researchers used nudges to test if users would change their profile. Most did not. In addition, while users were always able to change their profile, few did. While the authors claimed that this supports the accuracy of their profiles, this could also be interpreted as evidence of the inertia bias. We can also imagine someone who continued to receive notifications based on her original privacy profile and acted according to them even if it she determines that this profile is not the best fit for her. Like certain manifestations of notice fatigue, failing to act on this intention would result in a failure to externally self-realize. In addition, her resulting privacy decisions would likely not be consistent with her values, a violation of self-unification. As long as PPAs do not account for the inertia bias, they cannot be said to be fully self-realizing or self-unifying.

Lastly, there is the challenge at the level of privacy controls themselves. There are a plethora of different privacy notice designs with different degrees of privacy control (Utz et al., 2019). Smartphone PPAs can also only modify preferences using the smartphone operating system's (OS) available controls. These controls (e.g., access to Contacts, Camera, Location) have been previously shown to be insufficient for capturing the privacy concerns of users (Felt et al., 2012). In addition, Terpstra et al. (2019) have purposed that a lack of meaningful privacy controls in conjunction with positive friction lead to user frustration, possibly undermining positive friction benefits such as value reflection. This frustration can be understood in terms of failures to self-realization and self-unification. If a user forms an intention (privacy preference) but the control is not present, this would be a frustration of self-realization. Because she is not acting according to her values, this would also constitute a failure of self-unification. Due to the limited level of granularity available to them, our ability to realize our commitments and make choices consistent with our commitments is still stifled with PPA use.

3.3 In the Context of Selecting Smartphone Applications

These challenges present three major implications for designing a PPA-like system, a value-centered privacy assistant (VcPA), to help users select smartphone applications. Firstly, like current PPAs, users of a VcPA may fail to update and change their privacy profile due to the inertia bias, hindering self-realization and self-unification. Secondly, these dimensions could be further undermined due to the lack of granularity problem. While the challenges to self-realization and self-unification of current PPAs are a result of the lack of granularity of either (1) privacy notice design (online PPAs) or (2) the OS' privacy setting controls, a PPA-like system assisting with smartphone application selection is essentially a dichotomous decision: either download or not download. Not only does this minimize self-realization, but this will prevent a user from realizing her commitments in a manner consistent (unifying) with her values by introducing additional double bind situations. For example, consider a user deliberating whether to download a social media app such as Instagram. While she greatly values her social connections, she also values control over her life. She must decide whether to join social media and allow her data to be collected or to abstain, neither of which will be in full accordance with her values. Thirdly, the aforementioned double bind situation concerning efficacy and PPAs will also need to be considered when selecting an app; someone who values efficiency may still face instances of double binds if selective notifications are based on the app's data collection practices and her privacy preferences alone.

3.4 Suggested Modifications when Developing a Value-centered Smartphone Privacy Assistant

There are a few possible modifications to current PPA design that could overcome these remaining challenges and create a value-centered privacy assistant (VcPA) (Table 1).

To rectify the efficiency value challenge, VcPA profiles could instead be based on the user's personal values. By basing profiles on values rather than privacy preferences alone, the VcPA could be more accurately tuned to prevent efficiency-based double bind situations and maximize efficiency. User tests, however, will be required to determine in what way values intersect with app data collection preferences and how this intersection could be operationalized as VcPA profiles. In addition, these profiles will also need to (1) be an *accurate* reflection of a user's values; (2) have an accessible and *understandable description*; (3) clearly state *how* privacy preferences and values are utilized in profile creation and assignment; and (4) be able to be *changed* to a different profile if notices are unhelpful. All of these considerations must be accounted for when designing and testing a VcPA prototype in order to rectify the double bind challenge without sacrificing accuracy and transparency.²⁸

²⁸ To this aim, a prototype VcPA has been designed and has been tested with users. For more on the technical development of the VcPA and its testing environment, please see Carter et al. (2022). At the time of writing, the research team is analyzing participant feedback and will report the findings in a forthcoming paper.

Self-realization and self-unification as defined by 4DT are minimized in part because PPAs are limited by the lack of relevant controls, which remain relevant to VcPAs through a download-or-not-choice that may make acting consistent with their values not possible. This challenge could be overcome in a VcPA by not only informing the user when an application is requesting privacy settings *inconsistent* with their values, but by also suggesting alternative apps to quickly link them out to similar applications that better align with their commitments. This increases the likelihood that they will be able to find an app whose data collection practices match their values, thus (better) upholding self-realization and self-unification.²⁹

In order to be more fully self-realizing and self-unifying, however, a VcPA will also have to tackle the inertia bias. Agreeing with others that there is a need for “learning” (or, at least room for change) in the privacy notice process (Terpstra et al., 2019), this could be tackled by periodically (but not excessively) “mining” user goals, values, and preferences—that is, checking in with the user that the notices are still a good fit for them by making them aware, perhaps by using an “exploratory notice.” It will be critical, however, that this “mining” process be sufficiently random and dispersed as to *not* encourage user action one way or another. If a user makes a choice that is inconsistent with what they have resolved to do as a result of this exploratory notice, this would be inconsistent with external self-realization (and possibly self-unification if the action goes against her values). In the recommender system literature, researchers have empirically studied how to limit the effects recommendations have in encouraging certain behavior, suggesting that a process of “continuous random exploration” could be a way to combat these phenomena (Jiang et al., 2019). A similar approach to continuous random exploration could be utilized to time exploratory notices. User tests will be required, however, to determine the right balance between providing users with the ability to learn and explore through exploratory notices and inadvertently nudging them in a non-autonomous direction.

4 Final Thoughts

Here, I have proposed a new goal for privacy notices—to *create the space* for value-centered privacy decision-making. This understands data privacy decisions as, ideally, a reflection of user values. After linking this to autonomy, I then explored how

²⁹ An important caveat here is the case of social media applications such as Instagram, mentioned in Sect. 3.3 as presenting a possible double bind situation. A VcPA that recommends alternative applications will only be effective at addressing double bind and self-realization/unification concerns if such an alternative is available. This may not be the case for social media and messaging applications. Returning to the example user that values both social connectivity and control when deliberating whether to join Instagram, an alternative with similar social value *to her* is likely not available if most of her friends and family are on Instagram. As a VcPA can only work within the existing app ecosystem, including social media apps that essentially have “monopolies” on online socialization, they are unlikely to be as effective at resolving social media app double bind situations. However, I would argue that the potential for VcPAs to promote value-centered decisions within other app contexts still makes them worthy of exploration.

we can design for this space while remaining mindful of insights from behavioral psychology. To inform the design of a smartphone assistant that creates this space for users, I first examined PPA technology using a 4DT lens. While PPAs, with their use of selective notices, are partly successful at creating the space for value-centered choice, a number of concerns around the inertia bias, double bind choices, and lack of controls remain. Using these insights to inform notice timing, content, and basis, I lastly proposed a value-centered, smartphone privacy assistant (VcPA) to help users make more value-centered decisions at one privacy decision point: smartphone app choices.

There are, however, various concerns of this approach that still need to be addressed. One aspect of PPAs in general—whether value-centered or not—that has not fully been addressed here is that the collection of data itself for running such a system could be utilized to undermine a user’s autonomy. For example, in a value-centered system, information concerning a user’s personal values and how they relate to their app choices is quite sensitive in nature and could be utilized by targeted advertisers to nudge users towards products more effectively. I think this is, however, fundamentally a technical challenge than a foundational threat to the VcPA concept. It is an issue of data protection that could be addressed by ensuring that the data required to run the VcPA is held locally on the user’s device and able to interact with an app store without disclosing any information about the user (Carter, 2021). Local data storage, however, could take up too much space on a user’s phone and make a VcPA impractical. These challenges will have to be addressed during the technical development of the assistant, in addition to determining how to design VcPA profiles based on personal values and the ideal frequency for the exploratory notices.³⁰

In addition, empirical studies will be required to validate the conceptualization of data privacy decisions as a reflection of personal values and to assess whether the VcPA notices designed to promote them are effective. To this end, a mixed-methods study is currently in progress. The aims of this study are to determine if (1) a VcPA system is beneficial over traditional smartphone privacy notifications and controls and (2) if such a system assists users to make app choices more consistent with their values. More fundamentally, this study also aims to assess the role of values in privacy-decision making and, in particular, smartphone application selection. Answering these questions will be critical to determining the feasibility and desirability of a smartphone privacy assistant that enables value-centered privacy decision-making.

Finally, VcPAs should not be the only mechanism by which we govern data flows and manage privacy. While I have suggested here that a VcPA could increase value-centered privacy decisions by re-purposing notices, data privacy decisions have broader implications than the individual who makes them. Individual data flows can, for example, be grouped into profiles by “big data analytics,” creating a group of algorithmically-related individuals with interests that go beyond that of any one group member (Mittelstadt, 2017). While I would disagree that these collective properties of data privacy means that respect for autonomy—the normative

³⁰ See footnote 28.

basis for value-centered privacy notices—has become null and void, I recognize that an autonomy-exclusive solution does not accurately capture all the dimensions of data privacy. I also want to acknowledge genuine concerns about the effectiveness and desirability of our current privacy regulations. For example, initial empirical data suggests that smartphone data tracking on Android devices has not significantly changed following the introduction of the GDPR (Kollnig et al., 2021), greatly drawing into question its effectiveness. Therefore, the ultimate remedy to the data privacy challenges our age—if one exists—will require synergy between individual and collective, and regulatory and technological, solutions. It will require VcPAs and GDPRs working in harmony.

Acknowledgements The author (a PhD candidate at University of Galway) is primarily supervised by Dr. Heike Felzmann (Discipline of Philosophy, University of Galway) and Prof. Dr. Mathieu d'Aquin (LORIA, Université de Lorraine, France), both of whom provided feedback on the concepts described in this work and on the manuscript itself. Additionally, secondary supervisors Dr. Kathryn Cormican (Enterprise Research Center, University of Galway) and Dr. Dave Lewis (ADAPT Center, Trinity College Dublin) supported concept development by contributing to the author's general supervision. The author would also like to thank the reviewers for their valuable feedback and Dr. John Danaher (School of Law, University of Galway) for recommending Killmister's work to her.

Author Contribution The author confirms sole responsibility for this work, including conceptualization and manuscript preparation. Acknowledgement is made to other persons who provided feedback on this work but do not meet the criteria for authorship (as outlined by *Digital Society's* authorship guidelines).

Funding This work is being conducted with financial support from the Science Foundation Ireland Center for Research Training in Digitally-Enhanced Reality (d-real) under Grant No. 18/CRT/6224. In accordance with funding Open Access mandates, the author has applied a CC BY public copyright license to this Author Accepted Manuscript version.

Code Availability Not applicable.

Declarations

Ethics Approval Not applicable.

Informed Consent Not applicable.

Competing Interests The author declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allen, A. L. (2013). An ethical duty to protect one's own information privacy? *Faculty Scholarship at Penn Law*, 451.
- Almuhimedi, H., Schaub, F., Sadeh, N., Adjerid, I., Acquisti, A., Gluck, J., Cranor, L., & Agarwal, Y. (2015). Your location has been shared 5,398 times! A field study on mobile app privacy nudging. In K. Inkpen & W. Woo (Eds.), *Proceedings of the Conference on Human Factors in Computing Systems (CHI)* (pp. 787–796). ACM. <https://doi.org/10.1145/2702123.2702210>
- Apple. (2021). *Data privacy day at Apple: Improving transparency and empowering users*. Apple Newsroom. <https://www.apple.com/newsroom/2021/01/data-privacy-day-at-apple-improving-transparency-and-empowering-users/?afid=p239%7C10078&cid=aos-us-aff-ir>
- Arvanitis, A., Kalliris, K., & Kaminiotis, K. (2020). Are defaults supportive of autonomy? An examination of nudges under the lens of Self-Determination Theory. *The Social Science Journal*, 1–11. <https://doi.org/10.1016/j.soscij.2019.08.003>
- Beauchamp, T. L. (2011). Informed consent: Its history, meaning, and present challenges. *Cambridge Quarterly of Healthcare Ethics*, 20(4), 515–523. <https://doi.org/10.1017/S0963180111000259>
- Ben-Shahar, O., & Schneider, C. E. (2010). The failure of mandated disclosure. In *U of Chicago Law and Economics* (No. 516; Clin Working Paper, Vol. 159, Issue 3). <https://doi.org/10.2139/ssrn.1567284>
- Brandt, A. M. (1978). Racism and research: The case of the Tuskegee syphilis study. *The Hastings Center Report*, 8(6), 21. <https://doi.org/10.2307/3561468>
- Brignull, H. (n.d.). *Deceptive design*. Retrieved September 29, 2022, from <https://www.deceptive.design>
- Calo, M. R. (2012). Against notice skepticism in privacy (and elsewhere). *Notre Dame Law Review*, 59(2011), 1027–1072. <https://ssrn.com/abstract=1790144>
- Carlson, R. V., Boyd, K. M., & Webb, D. J. (2004). The revision of the Declaration of Helsinki: Past, present and future. *British Journal of Clinical Pharmacology*, 57(6), 695–713. <https://doi.org/10.1111/j.1365-2125.2004.02103.x>
- Carter, S. E. (2021). Is downloading this app consistent with my values? Conceptualizing a value-centered privacy assistant. In D. Dennehy, A. Griva, N. Pouloudi, Y. Dwivedi, I. Pappas, & M. Mäntymäki (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 12896 LNCS* (pp. 285–291). Springer International Publishing. https://doi.org/10.1007/978-3-030-85447-8_25
- Carter, S. E., Tiddi, I., & Spagnuolo, D. (2022). A "Mock App Store" interface for virtual privacy assistants. In S. Schlobach, M. Pérez-Ortiz, & M. Tielman (Eds.), *HAI2022: Augmenting human intellect: Proceeding of the first international conference on hybrid human-artificial intelligence: Vol. 354 of Frontiers in Artificial Intelligence and Applications* (pp. 266–268). IOS Press. <https://doi.org/10.3233/FAIA220212>
- Childress, J. F. (1990). The place of autonomy in bioethics. *The Hastings Center Report*, 20(1), 12–17. <https://doi.org/10.2307/3562967>
- Chitkara, S., Gothoskar, N., Harish, S., Hong, J., & Agarwal, Y. (2017). Does this app really need my location? Context-aware privacy management for smartphones. In S. Santini (Ed.), *Proceedings of the Interactive, Mobile, Wearable and Ubiquitous Technologies (UbiComp)* (Vol. 1, Issue 3, pp. 1–22). ACM. <https://doi.org/10.1145/3132029>
- Cohen, J.E. (2013). *What Privacy Is For*. Harvard Law Review. 126(7):1904–1933
- Cox, A. L., Gould, S., Cecchinato, M. E., Iacovides, I., & Renfree, I. (2016). Design frictions for mindful interactions: The case for microboundaries. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)* (pp.1389–1397). <https://doi.org/10.1145/2851581.2892410>
- Das, A., Degeling, M., Smullen, D., & Sadeh, N. (2018). Personalized privacy assistants for the internet of things: Providing users with notice and choice. *IEEE Pervasive Computing*, 17(3), 35–46. <https://doi.org/10.1109/MPRV.2018.03367733>
- Degeling, M., Utz, C., Lentzsch, C., Hosseini, H., Schaub, F., & Holz, T. (2019). We value your privacy ... Now take some cookies: Measuring the GDPR's impact on web privacy. In *Symposium on Network and Distributed System Security. NDSS*. <https://doi.org/10.14722/ndss.2019.23378>
- Dworkin, G. (1988a). Paternalism: Some second thoughts. In *The Theory and Practice of Autonomy* (pp. 121–129). Cambridge University Press. <https://doi.org/10.1017/cbo9780511625206.009>
- Dworkin, G. (1988b). The nature of autonomy. In *The Theory and Practice of Autonomy* (pp. 1–20). Cambridge University Press. <https://doi.org/10.3402/nstep.v1.28479>

- Regulation (EU) 2016/679. *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)*. European Parliament and Council. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679>
- Felt, A. P., Egelman, S., & Wagner, D. (2012). I've got 99 problems, but vibration ain't one: A survey of smartphone users' concerns. In W. Enck & X. Jiang (Eds.), *Workshop on Security and Privacy in Smartphones and Mobile Devices (SPSM)* (pp. 33–43). ACM. <https://doi.org/10.1145/2381934.2381943>
- Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy*, 68(1), 5–20. <https://doi.org/10.2307/j.ctvvh84xp.15>
- Friedman, B., & Nissenbaum, H. (1997). Software agents and user autonomy. In *Proceedings of the First International Conference on Autonomous Agents* (pp.466–469). ACM.
- Friedman, B. (1996). Value-sensitive design. *Interactions*, 3, 16–23. <https://doi.org/10.1145/242485.242493>
- Friedman, Batya, Kahn, P. H., & Borning, A. (2008). Value sensitive design and information systems. In K. E. Himma & H. T. Tavani (Eds.), *The Handbook of Information and Computer Ethics* (1st ed., pp. 69–103). Wiley. <https://doi.org/10.1002/9780470281819.ch4>
- Grassl, P., Schraffenberger, H., Zuiderveen Borgesius, F., & Buijzen, M. (2021). Dark and bright patterns in cookie consent requests. *Journal of Digital Social Research*, 3(1), 1–38. <https://doi.org/10.31234/osf.io/gqs5h>
- Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The dark (patterns) side of UX design. In A. Dey, E. Cutrell, & M. C. Shruemel (Eds.), *Proceedings of the Conference on Human Factors in Computing Systems (CHI)* (pp. 1–14). ACM. <https://doi.org/10.1145/3173574.3174108>
- Green, Y. (2022). Tech companies are reconsidering an old enemy. *WIRED*. <https://www.wired.com/story/friction-social-media-moderation/>
- Hausman, D. M., & Welch, B. (2010). Debate: To nudge or not to nudge. *Journal of Political Philosophy*, 18(1), 123–136. <https://doi.org/10.1111/j.1467-9760.2009.00351.x>
- Hendl, T., Chung, R., & Wild, V. (2020). Pandemic surveillance and racialized subpopulations: Mitigating vulnerabilities in COVID-19 apps. *Journal of Bioethical Inquiry*. <https://doi.org/10.1007/s11673-020-10034-7>
- Hoofnagle, C. J., & Urban, J. M. (2014). Alan Westin's privacy homo economicus. *Wake Forest Law Review*, 14(1), 261–351. <https://doi.org/10.31235/osf.io/ta2z3>
- Jensen, C., & Potts, C. (2004). Privacy policies as decision-making tools. In M. Tscheligi & E. Dykstra-Erickson (Eds.), *Proceedings of the Conference on Human Factors in Computing Systems (CHI)* (Vol. 6, Issue 1, pp. 471–478). <https://doi.org/10.1145/985692.985752>
- Jiang, R., Chiappa, S., Lattimore, T., Gyorgy, A., & Kohli, P. (2019). Degenerate feedback loops in recommender systems. In V. Coniter, G. Hadfield, & S. Vallor (Eds.), *Conference on Artificial Intelligence, Ethics, and Society (AIIES'19) - Proceedings* (pp. 383–390). ACM. <https://doi.org/10.1145/3306618.3314288>
- Jongsma, K. R., Bredenoord, A. L., & Lucivero, F. (2018). Digital medicine: An opportunity to revisit the role of bioethicists. *American Journal of Bioethics*, 18(9), 69–70. <https://doi.org/10.1080/15265161.2018.1498952>
- Kahneman, D. (2011). *Thinking, fast and slow*. Penguin Books.
- Kelley, P. G., Cranor, L. F., & Sadeh, N. (2013). Privacy as part of the app decision-making process. In S. Bødker, S. Brewster, P. Baudisch, M. Beaudouin-Lafon, & W. E. Mackay (Eds.), *Proceedings of the Conference on Human Factors in Computing Systems (CHI)* (pp. 3393–3402). ACM. <https://doi.org/10.1145/2470654.2466466>
- Killmister, S. (2017). *Taking the measure of autonomy: A four-dimensional theory of self-governance* (1st ed.). Routledge. <https://doi.org/10.4324/9781315204932>
- Kitchin, R. (2020). *Using digital technologies to tackle the spread of the coronavirus : Panacea or folly?* (No. 44; The Programmable City). <https://progcity.maynoothuniversity.ie/wp-content/uploads/2020/04/Digital-tech-spread-of-coronavirus-Rob-Kitchin-PC-WP44.pdf>
- Klugman, C. M., Dunn, L. B., Schwartz, J., & Cohen, I. G. (2018). The ethics of smart pills and self-acting devices: Autonomy, truth-telling, and trust at the dawn of digital medicine. *American Journal of Bioethics*, 18(9), 38–47. <https://doi.org/10.1080/15265161.2018.1498933>
- Kollnig, K., Binns, R., Kleek, M. Van, Lyngs, U., Zhao, J., Tinsman, C., & Shadbolt, N. (2021). Before and after GDPR: tracking in mobile apps. *Internet Policy Review*, 10(4).
- Lewis, M. (2017). *The Undoing Project*. Allen Lane.

- Liu, B., Andersen, M. S., Schaub, F., Almuhiemedi, H., Zhang, S., Sadeh, N., Acquisti, A., & Agarwal, Y. (2016). Follow my recommendations: A personalized privacy assistant for mobile app permissions. In M. E. Zurko, S. Consolvo, & M. Smith (Eds.), *Proceedings of the Twelfth Symposium on Usable Privacy and Security (SOUPS)* (pp. 27–41). USENIX.
- Liu, B., Lin, J., & Sadeh, N. (2014). Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help? In A. Broder, K. Shim, & T. Suel (Eds.), *Proceedings of the 23rd International Conference on World Wide Web* (pp. 201–211). ACM. <https://doi.org/10.1145/2566486.2568035>
- Luciano, F. (2020). Mind the app—Considerations on the ethical risks of COVID-19 apps. *Philosophy and Technology*, 33(2), 167–172. <https://doi.org/10.1007/s13347-020-00408-5>
- Lucivero, F., & Jongsma, K. R. (2018). A mobile revolution for healthcare? Setting the agenda for bioethics. *Journal of Medical Ethics*, 44(10), 685–689. <https://doi.org/10.1136/medethics-2017-104741>
- Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale: Findings from a crawl of 11 K shopping websites. In Er. Gilbert & K. Karahalios (Eds.), *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW)* (Vol. 3, pp. 81:1–81:32). ACM. <https://doi.org/10.1145/3359183>
- McDonald, A. M., & Cranor, L. F. (2008). The cost of reading privacy policies. *A Journal of Law and Policy for the Information Society*, 4(3), 543–568. <https://kb.osu.edu/handle/1811172839>
- Mejtoft, T., Hale, S., & Söderström, U. (2019). Design friction. *Proceedings of the 31st European Congress on Cognitive Ergonomics*, 41–44. <https://doi.org/10.1145/3335082.3335106>
- Mittelstadt, B. (2017). From individual to group privacy in big data analytics. *Philosophy and Technology*, 30(4), 475–494. <https://doi.org/10.1007/s13347-017-0253-7>
- Morley, J., Cows, J., Taddeo, M., & Floridi, L. (2020). Ethical guidelines for COVID-19 tracing apps. *Nature*, 582, 29–31.
- Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, 79(1), 119–158.
- O’Neill, O. (2002). *Autonomy and Trust in Bioethics*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511606250>
- O’Neill, O. (2020). Trust and accountability in a digital age. *Philosophy*, 95(1), 3–17. <https://doi.org/10.1017/S0031819119000457>
- Obar, J. A., & Oeldorf-Hirsch, A. (2018). The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, 1–20. <https://doi.org/10.2139/ssrn.2757465>
- Peters, D., Calvo, R. A., & Ryan, R. M. (2018). Designing for motivation, engagement and wellbeing in digital experience. *Frontiers in Psychology*, 9(MAY), 1–15. <https://doi.org/10.3389/fpsyg.2018.00797>
- Roessler, B., & Mokrosinska, D. (2013). Privacy and social interaction. *Philosophy and Social Criticism*, 39(8), 771–791. <https://doi.org/10.1177/0191453713494968>
- Ryan, K. J., Brady, J. V., Cooke, R. E., Height, D. I., Jonsen, A. R., King, P., Lebacqz, K., Louisell, D. W., Seldin, D. W., Stellar, E., & Turtle, R. H. (1979). *The Belmont Report*. Department of Health, Education, and Welfare (US). <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>
- Ryan, R. M., & Deci, E. L. (2004). Autonomy is no illusion. In J. Greenberg, S. L. Koole, & T. Pyszczynski (Eds.), *Handbook of Experimental Existential Psychology* (1st ed., pp. 455–485). The Guilford Press.
- Schaub, F., Balebako, R., Durity, A. L., & Cranor, L. F. (2015). A design space for effective privacy notices. In L. F. Cranor, B. Robert, & S. Consolvo (Eds.), *Proceedings of the 11th Symposium on Usable Privacy and Security (SOUPS 2015)* (pp. 1–17). USEBIX Association. <https://doi.org/10.1017/9781316831960.021>
- Schmidt, A. T., & Engelen, B. (2020). The ethics of nudging: An overview. *Philosophy Compass*, 15(4), 1–13. <https://doi.org/10.1111/phc3.12658>
- Solove, D. J. (2002). Conceptualizing privacy. *California Law Review*, 90(4), 1087–1155. <https://doi.org/10.2307/3481326>
- Solove, D. J. (2007). “I’ve got nothing to hide” and other misunderstandings of privacy. *San Diego Law Review*, 44, 745–773. <https://ssrn.com/abstract=998565>
- Solove, D. J. (2020). The myth of the privacy paradox. *GW Law School Public Law and Legal Theory Paper*. <https://doi.org/10.2139/ssrn.3536265>
- Spiekermann, S., Grossklags, J., & Berendt, B. (2001). E-privacy in 2nd generation E-commerce: Privacy preferences versus actual behavior. In M. P. Wellman & Y. Shoham (Eds.), *Proceedings of the Conference on Electronic Commerce (EC)* (pp. 38–47). ACM.
- Story, P., Cranor, L. F., Smullen, D., Sadeh, N., Acquisti, A., & Schaub, F. (2020). From intent to action: Nudging users towards secure mobile payments. In *Proceedings of the 16th Symposium on Usable Privacy and Security (SOUPS)* (pp. 379–416).

- Susser, D. (2019). Notice after notice-and-consent: Why privacy disclosures are valuable even if consent frameworks aren't. *Journal of Information Policy*, 9, 37–62. <https://doi.org/10.5325/jinfopoli.9.2019.0132>
- Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review*, 8(2), 1–22. <https://doi.org/10.14763/2019.2.1410>
- Terpstra, A., Schouten, A. P., Rooij, A. de, & Leenes, R. E. (2019). Improving privacy choice through design: How designing for reflection could support privacy self-management. *First Monday*, 24(6). <https://doi.org/10.5210/fm.v24i7.9358>
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Caravan Books.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- Utz, C., Degeling, M., Fahl, S., Schaub, F., & Holz, T. (2019). (Un)informed consent: Studying GDPR consent notices in the field. In D. M. Freeman, A. Mitrokotsa, & A. Sinha (Eds.), *Proceedings of the ACM Conference on Computer and Communications Security* (pp. 973–990). ACM. <https://doi.org/10.1145/3319535.3354212>
- Waldman, A. E. (2015). Privacy as trust: Sharing personal information in a networked world. *University of Miami Law Review*, 69(3), 559–630. <https://doi.org/10.2139/ssrn.2309632>
- Waldman, A. E. (2018). Privacy, Notice, and Design. *Stan Tech Law Review*. <https://doi.org/10.2139/ssrn.2780305>
- Warberg, L., Acquisti, A., & Sicker, D. (2019). Can privacy nudges be tailored to individuals' decision making and personality traits? In J. Domingo-Ferrer (Ed.), *Workshop on Privacy and Electronic Society (WPES)* (pp. 175–197). ACM. <https://doi.org/10.1145/3338498.3358656>
- World Medical Association. (2013). Declaration of Helsinki - Ethical principles for scientific requirements and research protocols. *Bulletin of the World Health Organization*, 79(4), 373. <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power* (1st ed.). New York: Profile Books.
- Zuiderveen Borgesius, F. J. (2015). Improving privacy protection in the area of behavioural targeting. *SSRN Electronic Journal*. <https://papers.ssrn.com/abstract=2654213>