Group Privacy: a Defence and an Interpretation

Family Name	Floridi
Given Name	Luciano
Division	Oxford Internet Institute
Organization	University of Oxford
Address	1 St Giles Oxford OX1 3JS United Kingdom
Email	luciano.floridi@oii.ox.ac.uk

Abstract

In this chapter I identify three problems affecting the plausibility of group privacy and argue in favour of their resolution. The first problem concerns the nature of the groups in question. I shall argue that groups are neither *discovered* nor *invented*, but *designed* by the level of abstraction (LoA) at which a specific analysis of a social system is developed. Their design is therefore justified insofar as the purpose, guiding the choice of the LoA, is justified. This should remove the objection that groups cannot have a right to privacy because groups are mere artefacts (there are no groups, only individuals) or that, even if there are groups, it is too difficult to deal with them. The second problem concerns the possibility of attributing rights to groups. I shall argue that the same logic of attribution of a right to individuals may be used to attribute a right to a group, provided one modifies the LoA and now treats the whole group itself as an individual. This should remove the objection that, even if groups exist and are manageable, they cannot be treated as holders of rights. The third problem concerns the possibility of attributing a right to privacy to groups. I shall argue that sometimes it is the group and only the group, not its members, that is correctly identified as the correct holder of a right to privacy. This should remove the objection that privacy, as a group right, is a right held not by a group as a group but rather by the group's members severally. The solutions of the three problems supports the thesis that an interpretation of privacy in terms of a protection of the information that constitutes an individual—both in terms of a single person and in terms of a group—is better suited than other interpretations to make sense of group privacy.

Introduction

The debate on Big Data (including Open Data) and Data Protection focuses on individual privacy. How can the latter be protected while taking advantage of the enormous potentialities offered by ever-larger data sets and ever-smarter algorithms and applications? The tension is sometimes presented as being asymmetric: between the *ethics* of privacy and the *politics* of security. In fact, it is ultimately ethical. Two moral duties need to be reconciled proactively: fostering human rights and improving human welfare. The tension is obvious if one considers medical contexts and biomedical big data, for example, where protection of patients' records and the cure or prevention of diseases need to go hand in hand.[1]

Currently, the balance between these two moral duties is implicitly understood within a classic ontological framework. The beneficiaries of the exercise of the two moral duties are the individual person vs. the whole society to which the individual belongs. At first sight, this may seem unproblematic. We work on the assumption that these are the only two 'weights' on the two sides of the scale. Such a framework is not mistaken, but it is dangerously reductive, and it should be expanded urgently. For there is a third 'weight' that must be taken into account by data protection: that of groups and their privacy.

The chapters in this volume provide a detailed analysis of the possibility of attributing a right to privacy to groups and sophisticated analyses of the scholarship behind the debate on group privacy, especially in modern legislation. In this contribution, I shall assume that it is *prima facie* plausible that groups may indeed enjoy such a right. However, there are at least three problems that may undermine such plausibility. I shall address them in the following pages with the hope that their solutions will facilitate the development of our ideas on group privacy.

The first problem, to be discussed in section one, concerns the nature of the groups in question. I shall argue that groups are neither *discovered* nor *invented*, but *designed* by the level of abstraction (LoA) at which a specific analysis of a social system is developed. Their design is therefore justified insofar as the purpose, guiding the choice of the LoA, is justified. This should remove the objection that groups cannot have a right to privacy because groups are mere artefacts (there are no groups, only individual persons to which groups are ultimately reducible) or that, even if there are groups, it is too difficult to deal with them.

---

[1] (Howe et al. 2008) and (Groves et al. 2013), for a review see (Mittelstadt and Floridi forthcoming-a). Most recent analyses of ethical problems in biomedical big data are provided in (Mittelstadt and Floridi forthcoming-b).

In section two, I shall address the next problem: assuming that there are groups and that they can be successfully managed, in what way can they be attributed rights? I shall argue that the same logic of attribution of a right to individual persons may be used to attribute a right to a group, provided one modifies the LoA and now treats the whole group as an individual in itself. I shall further argue that attributing a right to a person *or* to that person's group need not be incompatible alternatives, that is, the 'or' may be sometimes read as inclusive (as a logical 'and/or' or Latin *vel*, not necessarily always as an *aut aut*). This should remove the objection that, even if groups exist and are manageable, they cannot and should not be treated as holders of rights.

In section three, I shall then show in what sense groups may enjoy a right to privacy as groups. This should remove the objection that privacy, as a group right, is a right held not by a group as a group but rather by the group's members severally. Sometimes it is the group and only the group, not its members, that is correctly identified as the correct holder of a right to privacy. The analogy here is with the right of self-determination, which is held by a nation as a whole, not merely by its members severally.

The solutions of the three problems listed above lead to a final set of considerations, in section four, about the nature of privacy that may be enjoyed by a group. There I shall argue that an interpretation of privacy in terms of a protection of the information that constitutes an individual—both in terms of a single person and in terms of a group—is better suited than other current interpretations to make sense of group privacy.

To conclude, I shall argue that there are groups, designed by our ways of modelling interactions between agents and patients (senders and receivers of actions); that they can be and are manageable as holders of rights; and, in particular, that groups can be the primary holders of a right to privacy when this is constitutive of their identities. If I am correct, there is plenty of work for legislators to do. Let us see whether I am.

1. How can there be groups?

The debate about the nature of groups in philosophy of law and social science is strictly related to two other debates. One, in analytic philosophy (Beebee and Sabbarton-Leary 2010; Campbell, O'rourke, and Slater 2011), concerns natural kinds and whether there are 'natural'—as opposed to only arbitrary—ways of grouping objects, events, or beings on the basis of some shared, intrinsic properties (essentialism), such as chemical composition in the case of all objects made of gold (where gold is the natural kind). The other debate,

in philosophy of biology (Panchen 1992; Laporte 2004; Richards 2010; Oderberg 2013), concerns the nature of species, and more generally biological taxonomies, and addresses the question whether one or more criteria (such as reproductive isolation, or what it looks like, or indeed, genome) may be sufficient to categorise species, or decide whether an organism belongs to one species or another.

The similarities between the three debates are due to the fact that they are particular versions of the more fundamental and long-term debate between nominalism (there are only individuals, or tokens) and realism (there are also universals, or types). The nominalists and the realists tend to agree on the existence of individuals. They are both happy with Alice, her golden ring, and her puppy. They disagree on the existence of groups (Alice's family), natural kinds (golden objects), and species (Canis familiaris), and, in some cases, on the order of ontological priority (in various forms of Platonism universals not only exist but also precede, in terms of logical order, their instantiations). In short, they disagree on whether groups, natural kinds, and species may be only subjective and observer-dependent, or also objective and observer-independent.

Such ontological disagreement about what there is in reality and how it is organised in itself is possible because it presupposes a common epistemological framework, which enables the nominalist and the realist to avoid arguing at cross purposes. This is the view that knowledge can provide direct access to the intrinsic nature of its referents, i.e. what there is (or isn't) in the world in itself; the *noumenon*, to use a Kantian terminology. Interestingly, the further we move away from natural sciences and the closer we get to social or engineering ones, the easier it is to see this as a mistaken assumption, which leads to a false dichotomy. According to the nominalist, social groups (to restrict now the issue to our current concerns) are *invented*. According to the realist, they are *discovered*. The truth is that they are *designed*, that is, they are the outcome of the coming together of the world and the mind. To be more precise, they result from the choices we make of the observables we wish to focus on, for specific purposes, and from the constraining affordances (data) provided by the systems we are analysing. Thus, the position I wish to defend about the ontology of social groups is anti-essentialist but not anti-realist.[2] Let me illustrate it with an analogy.

Let us call a set of observables a level of abstraction (LoA). There is a LoA at which there are only individual buildings, and Alice's new flat and Bob's Victorian semi-detached

---

2 For a similar position in philosophy of biology see (Khalidi 2013).

house cannot possibly form a group. The two buildings may be regulated by very different kinds of legislation, provide different affordances, appeal to different home buyers, and so forth. They are so different from each other that they could never form a group. But then there is also a LoA at which both are two-bedroom accommodations in Oxford, for example, subject to the same local council taxation, perhaps rented from the same owner, and so forth. They are obviously part of a group. Asking whether a set of entities does or does not form a group independently of why one is asking the question in the first place, that is, independently of any interest in which features of the objects should count (e.g. the number of bedrooms for taxation purposes) is like asking the absolute price of a car without accepting any currency as a means to convey it.

There are of course groups that seem to us more natural. Yet the naturalness of a grouping is just a function of the intuitiveness of a LoA, that is, it is epistemological, not ontological. Referring to salad, tomatoes and potatoes as a group called food seems something as observer-independent and objective as possible, but this is only because we assume our own interests as organisms and eaters as the natural, intuitive, and relevant LoA. To a tiger, they would all look as unrelated and as eatable as grass and leaves to us. Accepting that our knowledge of the world is obtained through different LoAs is not to say that anything goes, and that the only alternative to nominalism and realism is some kind of untenable relativism. It is to say that absolute questions asked in a logical space lacking any references (LoA) and orientation (interest, purpose) are an absolute mess, and that *relationalism* (or *liminalism*, if you prefer a fancier word) is a better alternative. Using the previous example, asking whether something is food means adopting the right LoA at which it makes sense to ask whether a specific substance can be a nutrient for a specific organism. Food is a relational (not a relative) concept: it takes a LoA with two relata to define it, yet not every LoA is correct and some LoAs will be more correct than others.

All this means that we cannot be naïvely nominalist or realist about our ontology, especially when it comes to complex objects such as social groups. Imagine reality in itself as a sender of messages. Reality, understood as the Big Radio, broadcasts a very wide spectrum of signals. We, humans, are able to receive some of them directly, some others indirectly. For example, the visible spectrum is the portion of the electromagnetic spectrum that is detectable by the human eye and this is our most fundamental LoA when it comes to visual perception; we can see invisible radiant energy (for example, infrared, electromagnetic radiation with longer wavelengths than those of visible light) through technological mediations. Out of all those signals, we make sense of the sender itself. It

would be utterly naïve to think that the signals are a description of the sender, yet this does not mean that they are any less real. We only have to admit that the Big Radio is not sending selfies. With two other, different analogies, we cook with some ingredients (data from the world) but the dish we obtain (information) is not a copy of the ingredients. Or, we build with some materials (data in the world), but the house we obtain (information) is not a copy of the bricks we used. Human knowledge works in this *constructionist* (not *constructivist*, mind) way, it is not *mimetic*, it is *poietic*. Some parts of this *poiesis* are heavily constrained by the signals we receive. In the long run, we ask more questions to get more data, as Francis Bacon already suggested. We manipulate the data to see what further data can be obtained, and all this leads to scientific theories, which are our best ways of making sense of the constraining affordances (my preferred definition of data) provided by the realities we are studying. Some other parts of this interpretation are more flexible and malleable, i.e. the constraining affordances provide much more latitude, and well-informed, rational disagreement is more difficult to resolve (think of economic policies during a financial crisis). There is nothing relativistic or anti-realist in this, in the same sense in which there is nothing relativistic or anti-realist in the dish we cooked or the house we built. Humanity has taken advantage of the signals sent by the Big Radio increasingly well and this is why our knowledge works so successfully. The fact that we find some grouping very intuitive is part of such a successful story. But we do not need to embrace any naïve essentialism, or representational theory of knowledge, or a correspondentist theory of truth to make sense of groups. We should think about our knowledge of the world not in terms of painting it but in terms of engineering a model of it. Grouping is part of the successful strategy through which we make sense of reality.

What follows from the previous outline is that social groups should neither be conceived as mere conventions or artefacts (invented) nor assumed to exist before the interest in identifying them is specified (discovered). They are more or less correctly and successfully designed by our epistemological interests and practices *together* with the ontological constraining affordances provided by the world.

Let us now return to the nature of social groups. Any social system of n individuals can be organised into 2^n groups. For example, Alice, Bob and Carol would give rise to the following eight groups (subsets): {}, {Alice}, {Bob}, {Carol}, {Alice, Bob}, {Alice, Carol}, {Bob, Carol}, {Alice, Bob, Carol}. It is obvious that the power set of a set (the group containing all possible groupings) soon becomes unmanageable. At the same time, privileging only some groups as 'real' may seem to be arbitrary. Why should {Alice, Carol}

count, but not {Bob, Carol}? Because both Alice and Carol are female? But what if the criterion is having a rare disease, which Bob and Carol share, but not Alice? Clearly what matters is the LoA (in our example gender or health) at which the data we have (in our example, Alice's, Bob's, and Carol's)—the constraining affordances—are transformed (modelled) into information that ends up generating a group. The logical order is therefore: purpose (why grouping individuals in this way), LoA (how grouping individuals in this way), result (the obtained group). With an elementary example,[3] in a legal class action first comes the interest in dealing with a specific issue. This sets the observables (the LoA), e.g. some Electrolux dryers are alleged to "contain defects that can cause them to catch fire due to lint buildup". Given this LoA, one can then identify the group, that is, who is eligible "if you purchased certain freestanding clothes dryers between Jan. 1, 2002 and Dec. 31, 2011, you could be eligible for benefits from the Electrolux class action settlement". The LoA designs the group of eligible people. Asking whether the group is discovered objectively or invented subjectively *before* the interest and LoA are specified is not even incorrect, it is just missing the point entirely. Of course, some social groups simply self-determine their own nature, by adopting the purpose and LoA at which they wish to be identified.

All this is particularly relevant in the case of group privacy because it would be a mistake to think that first one has to establish the existence of a group, then the presence of a group's right to privacy, and then the potential infringement of that group's privacy through some Information and Communication Technologies (ICTs) application. If this were the case, we would be facing an intractable problem, because the identification of groups *a priori*, independently of the identification first of any interest or purpose (and hence LoA) that determines the grouping, is open to endless debate. Luckily, the process in practice is rather the opposite. First comes the interest (usually, but not necessarily pursued through the application of a technology) in clustering people in some groups. For example, a retailer may be interested in reaching all pregnant women in Oxford in order to advertise some products. This group may or may not overlap with other, pre-existing, intuitive groups, yet this does not matter (although this can be confusing when approaching the issue from a nominalist vs. realist perspective), even when the interested practices in question may be self-reflective, i.e. even when individuals may wish to identify themselves as members of a group, for this too is an epistemological choice (note that the

[3] See http://topclassactions.com/lawsuit-settlements/open-lawsuit-settlements/30306-electrolux-dryer-class-action-settlement/

mistake here would be to attempt to identify all possible social groups in Oxford and then check whether their rights have been infringed, an impossible task). Then comes the potential breach of the privacy of such a group as a group ('as a group' is an assumption that still needs to be defended below, bear with me). Note that what constitutes the group is also what makes group privacy possible. And finally comes the right of the group to see the situation redressed. In short, there is no nominalist objection to group privacy because it is the very same interested practices determining the grouping of people that also delineate the resulting groups as potential holders of a right to privacy, which then the group can exercise. Profiling is not a *descriptive* practice, it is a *designing* one, and it comes with the consequence of creating the condition of possibility of the profiled individuals, now constituted as a group by the very act of profiling, to act as a group in order to claim respect for its own privacy. Of course the grouped (profiled) individuals may not know that they have been profiled, e.g. by automatic algorithms, and may never discover that they have been treated as a group. This is not the point. What I am arguing is that if they end up being profiled and this profiling becomes explicit, what gives the group the initial possibility of reacting to it is the "interested" practice of profiling it in the first place, not some pre-existing ontological status of the group as a group, that would allegedly predate the profiling. With an analogy, the slice may not know that is has been severed from the rest of the cake, but if it realises that it has been it also realises that it was the severing it from the cake that gave rise to its identity, which did not precede the severing process itself. With one more analogy, grouping cuts both sides of the same piece of paper, the social (who is and is not in a group) and the ethical (which group has a right to privacy); you cannot have one without the other. All this explains why profiled individuals often object not so much to the treatment of themselves as members of a group but to the very profiling in the first place. It is not being a slice the problem, the problem is being severed from the cake in the first place.

The next question then becomes: if groups are constituted by the interested practices of grouping, for a purpose, and at a particular LoA, in what sense, if any, can they have a right?

2. How can a group have rights?

Groups are the social, qualitatively richer instance of mathematical sets. This is useful, because, by looking at sets, it is much easier to clarify in what sense a group and its

members may or may not share the same property, including a particular right. Let me explain.

Imagine a small departmental library. We need to move it from one building to another. We decide to move first all books with authors from A to D. Clearly the pile of books does not share that property, that is, it would be meaningless to ask whether the pile has an author. Next, suppose we are concerned about the fact that each of our books is inflammable. The concern remains once we realise that the pile inherits the same property. Third, we try to lift the pile and notice that it has now acquired a property that none of the books has: it is too heavy to be moved by a single person, despite the fact that each book in it is reasonably small and light. With a sigh, we finally wish books could fly from one building to another, but they do not, and neither do piles of them. This example illustrates the four possible cases in which sets and their members may or may not share a property (see Fig. 1). I introduced them in order of importance. The first case generates a common fallacy. The last case is not relevant to our discussion.

	has the property *F*			
	1	2	3	4
Members	Yes	Yes	No	No
Set	No	Yes	Yes	No
Example	Author	Inflammable	Heavy	Flies

Fig. 1 The relations of commonality of properties between sets and their members in the case of books and piles of books

The debate on whether groups (sets) may have rights (the property *F*) can be clarified by using the four columns in Fig. 1.[4] Sceptics subscribe to position 1: rights are properties that qualify only members of a group, not a group; speaking of a group right makes no sense and it is based on a fallacy. Moderate supporters of group rights tend to sit in the middle, subscribing to position 2: a group has rights, but only because each individual person constituting it has such rights. Finally, strong supporters of the idea of group rights subscribe to position 3: there are some kinds of rights that belong only to a group as a group, not to a group insofar as it is constituted by individual persons who enjoy those rights. In this case, it is important to understand that the group itself acts as an individual, to which a right is attributed. This is the case with political rights, as we have already seen: it is a shift in the LoA that allows one to consider a whole nation as having a right to self-determination as an individual agent. The point is important not only for the sake of clarity, but also because we saw that determining the LoA is what makes talking about groups ontologically unproblematic. By grouping people according to specific criteria we create an individual (the group), which can both be targeted and claim to have rights as a group.

The debate between the sceptical, the moderate and the strong position about group rights leads us to the last problem I wish to address here: how a group can have a right to privacy.

[4] For the sake of simplicity in what follows I shall assume that if members of a set and the set have the same property *F* this is because the set inherits *F* from its members. This is not necessarily the case and things become more complicated if we include the case in which both members and their set may have the same property *F* but for different reasons, that is, if the relation between the *F* of the members and the *F* of their set is not one of inheritance but of repeated occurrence. For example, the set of all books without an author is also without an author, but not because of them, but because authorship does not qualify sets of books, only books.

3. How can a group have a right to privacy?

One problem with privacy is that it is unclear whether, if it applies to groups, it may apply sometimes in the moderate and sometimes in the strong sense. Consider the following two cases.

A new California Privacy Law for Minors took effect as of January 1, 2015.[5] Entitled "Privacy Rights for California Minors in the Digital World", it gives minors the right to delete content that they posted to a website, social media profile, or online service while under the age of 18. It also includes restrictions on marketing or advertising some specified products and services to minors. This law seems a case of moderate group privacy. It is phrased in terms of protection of the individual person (the term "minor" is used, in line with Privacy Law, to mean natural person individual under the age of 18 who resides in California) and it seems obvious from the text that any reference to minors as a group (the "General Audience Property(ies)") is only a shortcut for a reference to each of its members. Minors have a right to see their personal information online erased only because each minor does. Talking of group privacy in this case is merely convenient but does not seem to add anything to our understanding of the phenomenon.

Consider next the case in which the close friends and relatives (the group) of a deceased person decide to hold a private funeral. Attendance is by invitation only, but this is not meant to make the funeral 'exclusive'. The desired privacy may be due to a need for intimacy, for respectful quietness, to protect grieving and reflection, or perhaps because of cultural or religious customs. Whatever the reasons, in this case it seems very counterintuitive to argue that each member of the group (each close friend or relative of the deceased) has a right to a private funeral, or that the privacy demanded is just the collection of all individual privacies. It seems more reasonable to admit that we are in the presence of a strong, social sense of group privacy. It is the whole group as a group that has a right to that specific kind of privacy.

If privacy applies to groups only in the moderate sense seen above (recall also the analogy with the pile of books, which is inflammable just because each book in it is), then there is interest in exploring its consequences, but not its nature. For if groups have a right to privacy only insofar as their members do, then all that can be said about moderate group privacy in terms of theory can also be said by reference to personal privacy – there is

---

[5] California S.B. 568 amends Division 8 of the California Business and Professions Code to add Chapter 22.1, see http://goo.gl/ODqtcO

nothing special in group privacy over and above all the personal privacies of the group members – yet this very reducibility also means that any defence of personal privacy must also take into account moderate group privacy, for affecting the latter does mean affecting the personal privacy of its members. I shall return to this point in the conclusion, where I will argue that even a moderate approach to group privacy requires taking the latter seriously in terms of legislation, in order to protect the privacy of the individual persons involved. If privacy applies to groups also in the strong sense seen above (recall also the analogy with the pile of books, which is heavy despite the fact that each book is light), then there is interest in exploring not only its consequences but also its nature, and this leads me to a final set of considerations.

4. What kind of privacy can group privacy be?

It is hard to elucidate the nature of group privacy—now understood in the strong sense clarified above—without a clear idea of what theory of privacy one is endorsing in the first place. Two theories are particularly popular in the current literature: the reductionist interpretation and the ownership-based interpretation. Neither is entirely satisfactory,[6] so I shall suggest a third one, based on the identity-constitutive nature of privacy, and argue that it is more suitable to understand strong group privacy.

The reductionist interpretation argues that the value of privacy rests on a variety of undesirable consequences that may be caused by its breach, either personally, such as distress, or socially, such as unfairness. Privacy is a utility, also in the sense of providing an essential condition of possibility of good human interactions, by preserving human dignity or by guaranteeing political checks and balances, for example.

The ownership-based interpretation argues that informational privacy needs to be respected because of each person's rights to bodily security and property, where 'property of *x*' is classically understood as the right to exclusive use of *x*. A person is said to own his or her information (information about him- or herself) and therefore to be entitled to control its whole life cycle, from generation to erasure through usage.

The two interpretations are not incompatible, but they stress different aspects of the value of privacy. The reductionist interpretation is more oriented towards a consequentialist assessment of privacy, in terms of cost–benefit analyses of its protection or violation. The ownership-based interpretation is more oriented towards a 'natural rights'

[6] See (Floridi 2013, 2014), for a detailed criticism, which is only summarized here insofar as it is relevant to the thesis defended in this chapter.

understanding of the value of privacy itself, in terms of private or intellectual property. Unsurprisingly, because they both belong to a pre-digital culture, they both compare privacy breach to physical trespass or unauthorised invasion of, or intrusion in, a metaphorical space or sphere of personal information, the accessibility and usage of which ought to be fully controlled by its owner and hence kept private. As I have argued elsewhere (Floridi 2013, 2014), neither interpretation is entirely satisfactory in many respects.

The reductionist interpretation defends the need for respect for privacy in view of the potential misuse of the information acquired. So it is certainly reasonable, especially from a consequentialist perspective, to extend it to groups. However, it seems to support at most a moderate interpretation of group privacy; and recall that this is interesting only in terms of consequences. If all we are arguing is that groups may enjoy some privacy only because their members do, any reference to group privacy is a mere shortcut. Furthermore, the reductionist interpretation may be inconsistent with pursuing and furthering social interests and welfare. Although it is obvious that some public personal information may need to be protected—especially against profiling or unrestrained electronic surveillance—it remains unclear, on a purely reductionist basis, whether a society devoid of any privacy may not be a better society after all, with a higher, common welfare. Indeed, it has been convincingly argued[7] that the defence of privacy in the home—that is, within that special group represented by a family—may actually be used as a subterfuge to hide the dark side of privacy: domestic abuse, neglect, or mistreatment. Precisely because of reductionist-only considerations, even in democratic societies we tend to acknowledge that the right to privacy can be overridden when other concerns and priorities, including public safety or national security, become more pressing. All this by putting some significant interpretative pressure on the "arbitrary" clause that qualifies article 12 of The Universal Declaration of Human Rights which states that

> No one shall be subjected to *arbitrary* [emphasis added] interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.

The ownership-based interpretation also falls short of being entirely satisfactory, for at least three reasons. First, informational contamination may undermine passive informational privacy. This is the unwilling acquisition of information or data, including

[7] See (Fineman and Mykitiuk 1994), and especially the chapter by Elizabeth M. Schneider 'The Violence of Privacy' a reprint of her article published in 1990.

mere noise, imposed on someone by some external source. Brainwashing may not occur often, but junk mail, or the case of a person chatting loudly on a phone nearby, are unfortunately common experiences of passive privacy breach, yet no informational ownership seems to be violated. Second, there is a problem of privacy in public contexts. Privacy—and especially group privacy, if there is such thing—is often exercised publicly, that is, in spaces that are socially, physically, and informationally shared: anyone can see what an individual person or group is doing downtown. How could a CCTV system be a breach of an individual's privacy if the individual in question is accessing a space that is public in all possible senses anyway? The ownership-based interpretation cannot provide a satisfactory answer. And finally, there is a metaphorical and imprecise use of the concept of 'information ownership', which cannot quite explain the lossless acquisition or usage of information. Information is not like a pizza: contrary to other things that one owns, one's personal information is not lost when acquired by someone else. Analyses of privacy based on 'ownership' of an 'informational space' are metaphorical twice over. All these difficulties make it less usable as a theory of group privacy. We need a better alternative, so here is a proposal.

Both the reductionist and the ownership-based interpretation fail to acknowledge the significant changes brought about by digital ICTs. They belong to an industrial culture of material goods, mechanical interactions, and of manufacturing/trading relations, so they rely on conceptual frameworks that are overstretched when trying to cope with the new challenges offered by an informational culture of services, networks, and usability. Interestingly, in their classic article The Right to Privacy, published in the Harvard Law Review in 1890, Samuel D. Warren and Louis Brandeis had already realised this limit with impressive insight:

> where the value of the production [of some information] is found not in the right to take the profits arising from publication, but in the peace of mind or the relief afforded by the ability to prevent any publication at all, *it is difficult to regard the right as one of property, in the common acceptation of the term* [emphasis added]. (Warren and Brandeis 1890), p. 25.

More than a century later, privacy requires a radical re-interpretation. Such a re-interpretation is achieved by considering each *individual person or group* as constituted by his, her or its information, and hence by understanding a breach of an individual's informational privacy as a form of aggression towards that individual's identity. This interpretation of privacy as having an identity-constituting value is consistent with the fact that ICTs can both erode and reinforce informational privacy, and hence that a positive effort needs to be made in order to support not only Privacy Enhancing Technologies but

also constructive applications, which may allow users to design, shape, and maintain their identities as informational agents. The value of privacy is both to be defended and enhanced.

The information flow needs some friction in order to keep firm the distinction between the macro multi-agent system (the society) and the identity of the micro multi-agent systems (the individual persons and groups) within it. Any society (even a utopian one) in which no informational privacy is possible is one in which no identity-constituting process can take place, no personal or group identity can be developed and maintained, and hence no welfare can be achieved, social welfare being only the sum of the individuals involved. The total 'transparency' of the infosphere that may be advocated by some reductionists achieves the protection of society only by erasing all identity and individuality, a 'final solution' for sure, but hardly one that the individuals themselves, constituting the society so protected, would be happy to embrace. The advantage of the identity-constituting interpretation of privacy over the reductionist one is that consequentialist concerns may override respect for privacy, whereas the identity-constituting interpretation, by equating its protection to the protection of individual identity, considers it a fundamental right. By default, the presumption should always be in favour of its respect, even when we admit that privacy may be negotiable to some degree in special circumstances.

Looking at the nature of an individual person or group as being constituted by that individual's information enables one to understand the right to privacy as a right to immunity from unknown, undesired, or unintentional changes in one's own identity as an informational entity, both actively and passively. Actively, because collecting, storing, reproducing, manipulating etc. Alice's or her family's information amounts now to stages in cloning and fixing (profiling) their identities. Passively, because breaching Alice's or her family's privacy may now consist in forcing the individual or her group to acquire unwanted information, thus altering their nature as informational entities without consent. The first difficulty facing the ownership-based interpretation is thus avoided.

The identity-constituting interpretation suggests that a group's informational sphere and the identity of a group are co-referential, or two sides of the same coin. The right to privacy, both in the active and in the passive sense just seen, shields the group's identity. This is why privacy is extremely valuable and ought to be respected. The second problem affecting the ownership-based interpretation is therefore also solved because violations of informational privacy are now more fruitfully compared to kidnapping rather

than trespassing. The advantage, in this change of perspective, is that it becomes possible to dispose of the false dichotomy qualifying privacy in public or in private contexts. Some information constitutes a group context-independently, and therefore a group is perfectly justified in wishing to preserve its integrity and uniqueness even in entirely public places. Trespassing makes no sense in a public space, but kidnapping (even of a whole group) is a crime independently of where it is committed.

As for the third problem, one may still argue that an individual group 'owns' its information, yet no longer in the metaphorical sense seen above, but in the precise sense in which a group is its information. 'Its' in 'its information' is not the same 'its' as in 'its land' but rather the same 'its' as in 'its memories', 'its culture', 'its choices', 'its rites and customs', and so forth. It expresses a sense of constitutive belonging, not of external ownership, a sense in which its information is part *of* it but is not a (legal) possession *by* it. Once again, it is worth quoting Warren and Brandeis, this time at length, even if they had in mind the individual person, rather than an individual group:

> [...] the protection afforded to thoughts, sentiments, and emotions [...] is merely an instance of the enforcement of the more general right of the individual to be let alone. It is like the right not to be assaulted or beaten, the right not to be imprisoned, the right not to be maliciously persecuted, the right not to be defamed [or, the right not to be kidnapped, my addition]. In each of these rights [...] there inheres the quality of being owned or possessed and [...] *there may be some propriety in speaking of those rights as property. But, obviously, they bear little resemblance to what is ordinarily comprehended under that term*. The principle [...] is in reality not the principle of private propriety but that of *inviolate personality* [emphasis added]. [...] the right to privacy, as part of the more general right to the immunity of the person, [is] *the right to one's personality* [emphasis added].

This identity-constituting conception of privacy and its value has started being appreciated by more mature, information societies, where the identity-constituting interpretation reshapes some of the assumptions behind a still 'industrial', 'modern', or 'Newtonian' conception of privacy. The following considerations illustrate such a transition.

If some information is finally acknowledged to be a constitutive part of personal and group identity, then one day it may become strictly illegal to trade in some kinds of information, exactly as it is illegal to trade in human organs (including one's own) or slaves. At the same time, we might relax our attitude towards some kinds of 'dead individual information' that, like 'dead pieces of oneself', are not really, or no longer, constitutive of a person or a group. Legally, Alice may not sell her kidney, but she may sell her hair or be rewarded for giving blood. Likewise, her family may not sell its members, even if they all, unanimously, accept such a practice, but it may sell the properties of one of its deceased members as a group.

We are constantly leaving behind a trail of data, pretty much in the same sense in which we are shedding a huge trail of dead cells. The fact that nowadays ICTs allow our data trails to be recorded, monitored, processed and used for social, political or commercial purposes is a strong reminder of our informational nature as individual persons and groups. It might be seen as a new level of environmentalism, as an increase in what is recycled and a decrease in what is wasted (not unlike what bacteria do with DNA available in the environment). At the moment, all this is just speculation and in the future it will probably be a matter of fine adjustments of ethical sensibilities, but the third Geneva Convention (1949) already provides a clear test of what might be considered 'dead personal information'. A prisoner of war need only give his or her name, rank, date of birth, and serial number and no form of coercion may be inflicted on him or her to secure any further information, of any kind. Even if we were all treated fairly as 'prisoners of the information society', our privacy would be well protected and yet there would still be some personal data that would be perfectly fine to share with any other agent, even hostile ones. It is not a binary question of all or nothing, but an analogue one of fine balance and degree.

A further issue that might be illuminated by looking at privacy from an identity-constituting perspective are those of confidentiality and intimacy, two intrinsically group-based phenomena. The sharing of private information with someone, implicitly (especially by doing things together), or explicitly, through communication, is based on a relation of profound trust that binds the people involved intimately. This coupling is achieved by allowing persons to be partly constituted as selves by the same information. The union of the persons in question forms a single unity, a supra-agent, or a new multi-agent individual, the group. Precisely because entering into a new supra-agent is a delicate and risky operation, care should be exercised before 'melding' oneself with other individuals by sharing personal information or its source, such as common experiences. This is the way I interpret the concluding sentence of *The Catcher in the Rye*, the famous novel by J. D. Salinger:

> Don't tell anybody anything. If you do, you start missing everybody. (Salinger 1951)

Confidentiality and intimacy create a bond that is hard and slow to forge properly, yet resilient to many external forces when finally in place, as the group (the supra-agent) is stronger than the constitutive agents themselves. Relatives, friends, classmates, fellows, colleagues, comrades, companions, partners, teammates, spouses and so forth may all have experienced the nature of such a bond, the stronger taste of a 'we'. But it is also a bond that is brittle and difficult to restore when it comes to internal betrayal, since the disclosure, deliberate or unintentional, of some personal information in violation of confidence can

entirely and irrecoverably destroy the intimacy and privacy of the new, supra-agent born out of the joining agents, by painfully introducing discord. The 'we' is strongly armoured against 'the other', but extremely fragile against internal betrayal by 'one of us'.

A final issue can be touched upon rather briefly: the identity-constituting interpretation stresses that privacy is also a matter of construction of an individual's own identity. The right to be left alone is also the right to be allowed to experiment with one's own life, to start again, without having records that mummify one's personal identity forever, taking away from the individual person or group the power to form and mould who or what the individual is and can be. Every day, an individual person or group may wish to build a different, possibly better, 'I' or 'we'. We never stop becoming ourselves, so protecting persons and group privacy also means allowing that person and group the freedom to construct and change herself or itself profoundly. The right to privacy is also the right to a live, renewable identity that one can shape freely. This is why it matters.

Conclusion

The idea that groups may have (at least something akin to) a right to privacy is not new (see for example (Bloustein 1978, 2003)) and it is open to debate (Bisaz 2012). But it has not received the attention it deserves, although the issue is becoming increasingly important. And this because, by far, ICTs treat most people not as individuals but as members of specific groups (or classes, collections, crowds, populations and their segments etc.), where the groups are the really interesting focus, as carriers of rights, values, and potential risks. Think of the owners of such and such kind of car, shoppers of such and such kinds of goods, people who like this type of music, or people who go to that sort of restaurant, cat owners, dog owners, people who live in a specific postal code, carriers of a specific gene, people affected by a particular disease, team fans … Especially big data is more likely to treat types (of customers, users, citizens, demographic population, etc.) rather than tokens (you, Alice, me…), and hence groups rather than individuals. But re-identifiable groups are *ipso facto* targetable groups. And membership in a sufficient number of groups can easily lead to the re-identification of individuals. Indeed, in terms of logic, two sets (even if they are infinite) are already sufficient to identify a singleton (a set with exactly one element). As an elementary example, suppose A is the infinite set of all integers including and larger than 1, and B is the infinite set of all integers including and smaller than 1, their intersection contains exactly one element, namely 1 ($A \cap B = 1$). It is therefore a very dangerous fallacy to think that, if we protect personal data that identify people

individually, the protection of groups of people will take care of itself. I have argued above that we should consider group privacy as something that is sometimes reducible to the individual privacy of its members, and sometimes as something that belongs to the group as a group. I have defended the plausibility of both moderate and strong group privacy. But I have also stressed that defending moderate group privacy is already crucial, in terms of the significant nature of its consequences. This is not the current view. In particular, a 'nominalist' approach (or informational ontology (Floridi 2003)) to group privacy—take care of each member separately and the group will automatically be fine too—is currently at the roots of European legislation. This defines a "Data Subject" as:

> An identified or identifiable person to whom specific personal data relates. It is someone who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more specific factors (physical, physiological, mental, economic, cultural, social). (European commission).

As a consequence, both the 1995 Directive and the new Regulation under discussion focus on individual persons. The philosophy informing the approach may be grasped by looking at the following recitals (emphases added):

> Whereas the principles of protection must apply to any information concerning an identified or identifiable *person*; whereas, to determine whether a *person* is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said *person*; whereas the principles of protection shall not apply to data rendered anonymous in such a way that the *data subject* is no longer identifiable […]. (Directive 95/46/Ec)

and, even more restrictively (notice the "natural"):

> The principles of protection should apply to any information concerning an identified or identifiable *natural* person. To determine whether a *natural* person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the individual. The principles of data protection should not apply to data rendered anonymous in such a way that the data subject is no longer identifiable. (Com(2012) 10 final 2012/0010 (Cod)).

Yet even from a nominalist perspective, we should acknowledge that both friendly and hostile users of big data may not care about Alice at all, but only about the fact whether Alice, whoever she is, belongs to the group that regularly goes to the local church, or mosque, or synagogue, uses Grindr, or has gone to a hospital licensed to carry out abortions, or indeed shares a feature of your choice. In military terminology, Alice is hardly ever a High Value Target, like a special and unique building. She is usually part of a High Pay-off Target, like a tank in a column of tanks. It is the column that matters.

As I have argued elsewhere (Floridi 2013) our current ethical approach is too anthropocentric (only natural persons count) and nominalist (only the single individual

person counts). We should take other kinds of individuals, including groups, into account. We need to be more inclusive because we are underestimating the risks involved in opening anonymised personal data to public use, in cases in which *groups* of people may still be easily identified and targeted. Such inclusiveness should not be too hard to achieve. After all, we already accept as ordinary the fact that groups as *agents* may infringe on someone's privacy. In the United States, we are used to considering as normal collective lawsuits (class actions) in which a group may sue a person or another group. And in Europe, consumer organisations regularly bring claims on behalf of the groups they represent. Clearly, there are cases in which the protection of a right requires a balance between the agents, issuing the action, and the patients, receiving the action.

There are very few Moby-Dicks. Most of us are sardines. The individual sardine may believe that the encircling net is trying to catch it. It is not. It is trying to catch the whole shoal. It is therefore the shoal that needs to be protected, if the sardine is to be saved. An ethics addressing each of us as if we were all special Moby-Dicks may be flattering and perhaps, in other respects, not entirely mistaken, but needs to be urgently upgraded. Sometimes the only way to protect a person is to protect the group to which that person belongs. Preferably before any disaster happens. This moderate sense of group privacy is the least we should begin to consider, as a first step towards a full recognition of strong group privacy.

Acknowledgements

I am very grateful to Massimo Durante for his most valuable comments on a previous draft of this article; to Ugo Pagallo and the participants in the panel "Open Data and Data Protection: Problems and Perspectives"—organised at the conference *Computers, Privacy and Data Protection 2014 (CPDP 2014) Reforming Data Protection: The Global Perspective*—for their feedback; to the participants in the workshop on group privacy held in Amsterdam on the 8th of September 2014 for valuable discussions; to David Sutcliffe for his skilful copyediting of the final version and many insightful comments that improved it significantly; and to Linnet Taylor for her feedback on a penultimate draft of this chapter, and some enlightening conversations on the topic of group privacy. Her chapter in this volume makes a strong and convincing case for the protection of group privacy in contexts of geolocated data, especially in LMIC (see also (Taylor forthcoming)).

References

Beebee, H., and Sabbarton-Leary, N. (2010). The Semantics and Metaphysics of Natural Kinds, Routledge Studies in Metaphysics. New York; London: Routledge.

Bisaz, C. (2012). The Concept of Group Rights in International Law: Groups as Contested Right-Holders, Subjects and Legal Persons. Leiden: Brill, Nijhoff.

Bloustein, E.d J. (1978). Individual and Group Privacy. New Brunswick, N.J.: Transaction Publishers.

Bloustein, E. J. (2003). Individual & Group Privacy. 2nd ed. New Brunswick, N.J.: Transaction Publishers.

Campbell, J. K., O'Rourke, M., Slater, M. H. (2011). *Carving Nature at Its Joints : Natural Kinds in Metaphysics and Science*, *Topics in Contemporary Philosophy*. Cambridge, Mass.; London: MIT Press.

COM(2012) 10 Final 2012/0010 (COD). Proposal for a Directive of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data by Competent Authorities for the Purposes of Prevention, Investigation, Detection or Prosecution of Criminal Offences or the Execution of Criminal Penalties, and the Free Movement of Such Data /* Com/2012/010 Final - 2012/0010 (Cod) */

Directive 95/46/EC. Directive 95/46/Ec of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data.

European Commission. Justice, Data Protection, Glossary, Data Subject: Http://Ec.Europa.Eu/Justice/Data-Protection/Glossary/Index_En.Htm.

Fineman, M. A., Mykitiuk, R. (Eds.) (1994). *The Public Nature of Private Violence: The Discovery of Domestic Abuse*. New York; London: Routledge.

Floridi, L. (2003). Informational Realism. In *Selected Papers from Conference on Computers and Philosophy-Volume 37*, John Weckert, J., Al-Saggaf, Y. (Eds.) 7-12. Australian Computer Society.

Floridi, L. (2013). *The Ethics of Information*. Oxford: Oxford University Press.

Floridi, L. (2014). *The Fourth Revolution - How the Infosphere Is Reshaping Human Reality*. Oxford: Oxford University Press.

Groves, P., Kayyali, B., Knott, D., van Kuiken, S. (2013). The 'Big Data' Revolution in Healthcare. *McKinsey Quarterly*.

Howe, D., Costanzo, M., Fey, P., Gojobori, T., Hannick, L., Hide, W., Hill, D. P., Kania, R., Schaeffer, M., St Pierre,S. (2008). Big Data: The Future of Biocuration. *Nature 455*(7209): 47-50.

Khalidi, M. A. (2013). *Natural Categories and Human Kinds: Classification in the Natural and Social Sciences*. Cambridge: Cambridge University Press.

LaPorte, J. (2004). *Natural Kinds and Conceptual Change*. Cambridge: Cambridge University Press.

Mittelstadt, B., Floridi, L. forthcoming-a. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Science and Engineering Ethics*.

Mittelstadt, B., Floridi, L. (Eds.) forthcoming-b. *The Ethics of Biomedical Big Data*, *Law, Governance and Technology*. New York: Springer.

Oderberg, D. S. (2013). *Classifying Reality*. Chichester: Wiley-Blackwell.

Panchen, A. L. (1992). *Classification, Evolution, and the Nature of Biology*. Cambridge: Cambridge University Press.

Richards, R. A. (2010). *The Species Problem: a Philosophical Analysis*. Cambridge: Cambridge University Press.

Salinger, J. D. (1951). *The Catcher in the Rye*. London: H. Hamilton.

Taylor, L. forthcoming. No Place to Hide? The Ethics and Analytics of Tracking Mobility Using African Mobile Phone Data.Online version available at http://www.academia.edu/4785050/No_place_to_hide_The_ethics_and_analytics_of_tracking_mobility_using_African_mobile_phone_data.

Warren, S., and Brandeis, L. D. (1890). The Right to Privacy. *Harvard Law Review 193*(4).