Philos. Technol. (2017) 30:475–494
DOI 10.1007/s13347-017-0253-7

RESEARCH ARTICLE

From Individual to Group Privacy in Big Data Analytics

Brent Mittelstadt[1]

Received: 18 September 2016 / Accepted: 13 January 2017 / Published online: 11 February 2017

Abstract Mature information societies are characterised by mass production of data that provide insight into human behaviour. Analytics (as in *big data analytics*) has arisen as a practice to make sense of the data trails generated through interactions with networked devices, platforms and organisations. Persistent knowledge describing the behaviours and characteristics of people can be constructed over time, linking individuals into groups or classes of interest to the platform. Analytics allows for a new type of algorithmically assembled group to be formed that does not necessarily align with classes or attributes already protected by privacy and anti-discrimination law or addressed in fairness- and discrimination-aware analytics. Individuals are linked according to offline identifiers (e.g. age, ethnicity, geographical location) and shared behavioural identity tokens, allowing for predictions and decisions to be taken at a group rather than individual level. This article examines the ethical significance of such *ad hoc* groups in analytics and argues that the privacy interests of algorithmically assembled groups in *inviolate personality* must be recognised alongside individual privacy rights. Algorithmically grouped individuals have a collective interest in the creation of information about the group, and actions taken on its behalf. Group privacy is proposed as a third interest to balance alongside individual privacy and social, commercial and epistemic benefits when assessing the ethical acceptability of analytics platforms.

Keywords Privacy · Big data · Information ethics · Data protection · Profiling · Computer ethics · Biomedicine · Medical research · Analytics · Group privacy · Automated decision-making · Discrimination detection

✉ Brent Mittelstadt
brent.mittelstadt@oii.ox.ac.uk

[1] Oxford Internet Institute, University of Oxford, 1 St. Giles, Oxford OX1 3JS, UK

1 Introduction[1]

Mature information societies are characterised by mass production of data that provide insight into human behaviour (Floridi 2016b). Data analytics has arisen to make sense of the data trails generated through interactions with networked devices, platforms and organisations (Burrell 2016; Grindrod 2014). Users are understood through small patterns or correlations with others in the system, by which individuals are clustered into meaningful groups (or classes) according to their behaviour, preferences and other characteristics[2] (Schermer 2011; Hildebrandt 2008). Analytics informs immediate responses to the needs and preferences of users as well as longer-term strategic planning and development by service and platform providers (Grindrod 2014). Decision-making in areas such as risk stratification, credit scoring, search and media filtration, market segmentation, employment, policing and criminal sentencing is now routinely informed by analytics.

Analytics allows persistent knowledge describing the behaviours and characteristics of people to be constructed over time, forming individuals into meaningful groups or classes. Individuals are linked according to offline identifiers (e.g. age, ethnicity, geographical location) and new behavioural identity tokens, allowing for predictions and decisions to be taken at a group rather than individual level (Grindrod 2014). These digital collective identifiers disrupt the long-standing link between the individual, identity and privacy.

Similar to individual privacy interests, algorithmically grouped individuals have a collective interest in how information describing the group is generated and used. However, equivalent privacy rights and duties do not exist for such *ad hoc* groups created by analytics. In this article, I examine the feasibility of granting algorithmically assembled groups a right to privacy. The article does not define a general theory of group privacy but instead examines the case for granting a specific informational privacy right—the right to inviolate personality—to ad hoc groups. This right would stand in contrast (but not opposition) to protections afforded to identifiable individuals in privacy and data protection law.

In international law, groups are defined by a shared background (e.g. culture or collective purpose). Such purposeful collectives are granted rights, including the right to assemble (e.g. Bloustein 1976; Bisaz 2012). In contrast, I argue that algorithmically constructed ad hoc groups lacking a shared background should also have their interests in privacy formally recognised by being granted a limited right to manage the group's identity. Group privacy is proposed as a third interest to balance with individual privacy rights and the social, commercial and epistemic benefits of analytics. While privacy-, fairness- and discrimination-aware analytics techniques have gone some way to protect the interests of such groups, a group privacy right would demand the scope of such techniques and legal privacy protections be expanded beyond currently protected 'offline' classes (and proxies thereof).

[1] I would like to acknowledge the extremely helpful feedback received from the reviewers at *Philosophy & Technology* in preparing this paper for publication.

[2] 'Big data' may be the wrong term to describe such data. Correlations can be found between "small amounts of densely connected metadata" that are not 'big' in terms of volume (Ananny 2016, 101).

The paper is structured as follows: Section 2 describes the need for group privacy rights in analytics, including how a group identity is formed that attributes but is not reducible to the individual identities of members. Section 3 then describes group privacy as a right to *inviolate personality*. Theoretical challenges to this right are then considered in Section 4, including an examination of what it means to be a rights-holder and the case for ad hoc groups to be considered rights-holders in parallel to the privacy rights of individuals. Section 5 then examines two models for implementing group privacy protections, noting challenges for both. Section 6 concludes with reflections on future work required to develop regulatory and technical protections for group privacy interests.

2 Ethical Significance of Groups and Identity in Analytics

This paper addresses analytics systems involving algorithmic classification or grouping of individuals to drive decision-making. Profiling algorithms identify correlations and make predictions about behaviour at a group level, albeit with groups (or profiles) that are constantly changing and redefined by the algorithm (Zarsky 2013). A simplistic example of such a group is "dog owners living in Wales aged 38–40 that exercise regularly." Being identified as a member of this group could drive a variety of automated decision-making with harmful or beneficial effects for individual members, such as a preferential rate for health insurance.

It can be argued that analytics does not create such groups per se but rather assigns weights or probabilities to inputs that represent individual data subjects. However, grouping of individuals is a core and unavoidable technique in algorithmic classification. Grouping can occur in two senses: either in the description of subjects or actions taken on the basis of probabilities and predictive analytics. For the former, data subjects must always be considered along a limited set of dimensions. Classifications are defined along these dimensions. A very simplistic example will suffice: assume we want to classify festival goers according to the festivals they have already attended (perhaps for the sake of predicting which festivals they will attend in the future). Assume we are considering two previous festivals, *P* and *Q*. Four possible groups of festival goers emerge according to these dimensions: *PQ* (attended P and Q), *PnQ* (attended P but not Q), *nPQ* (attended Q but not P), or *nPnQ* (attended neither P nor Q). While a certain group may remain temporarily empty, for instance if we lack data on any festival goers that have not attended either festival in question, the possible groupings are limited according to the number of dimensions considered.

This logic applies to descriptive and predictive classifications; the number of possible groups will always be derived from the number of dimensions under consideration. For predictive analytics, dimensions can also refer to non-descriptive choices, for instance the choice to deliver a certain advertisement due to observations of prior actions. In this case, each data subject will be understood through a series of probabilities in addition to observations; for instance, based on your attendance of festivals P and Q and your expressed musical interests, we predict with 75% confidence that you will attend festival *R*. Some may wish to argue that such probabilistic reasoning is not a type of grouping. However, probabilities are calculated to drive decision-making or, at a minimum, learn something about the data subject, i.e. to classify her for future

decision-making. At the point, an action is taken on the basis of a probability, a group is formed of all data subjects receiving the action, even if certainty is lacking that the grouping is accurate. Once again, the possible range of actions, and thus groups, will be limited by the dimensions under consideration.

Algorithmic classification can be considered a privacy invasive practice that is not yet sufficiently regulated due to a prevailing regulatory focus on identifiable individuals. Data must normally contain identifiers (e.g. name, address) to be legally classified as *personal* and thus warrant stronger limits on processing and exchange (cf. European Commission 2012). An individual's capacity to manage data about herself ends once identifiers are irretrievably removed.[3] The individualistic focus of existing European privacy protections incorrectly suggests that privacy cannot be violated without identifiability.

Analytics raises a unique challenge for existing theories of privacy by creating groups defined by identity tokens not reducible to, or owned by, individual members of the group. In turn, traditional identifiers (e.g. name, address) are increasingly irrelevant in analytics to learn something about people. Individuals can be clustered according to behaviours, preferences and other characteristics without being identified (van der Sloot 2014). Such methods find individuals interesting only to the extent that they can be correlated with others (Floridi 2014; Vedder 1999); advertisers care not who they are advertising to but whether their advertisements reach a target market. Members of the group (e.g. a market segment) need not be *identified* but rather *classified* to be effectively targeted. Connections between individuals revealing particular risks or behaviours are of interest, not the identifiable individual as such (Floridi 2014; Vedder 1999). Algorithmic classification tells us something about individuals through their associations with ad hoc, ephemeral but insightful groups of allegedly similar people.

Identifiability operates at two levels here. The individual's offline identity consists of the types of identifiers currently protected under data protection law (e.g. name, address, national insurance number, unique identifier). Offline identities always describe one and only one unique person; they are a way to ensure a persistent record can be assembled over time. The offline identity of the individual is irrelevant to the formation of an ad hoc group and learning things about it.

In analytics systems, individuals also possess a profiling identity constructed from connections with groups of other data subjects based upon dimensions (e.g. behaviours, demographic attributes) deemed relevant (see Fig. 1). An individual has a profiling identity consisting of all the ad hoc groups into which she is placed (see Fig. 2). The features that define these groups are non-random behaviours and attributes, referred to here as 'behavioural identity tokens' (cf. Grindrod 2014). Similar to some shared offline identifiers (e.g. culture, ethnicity, post code), ownership of behavioural identity tokens is distributed across members of the group (see Fig. 3). Actions by members can change the tokens; a member defaulting on a loan can, for example, increase the perceived risk of future loans to members. These changes, and decisions based upon

[3] This is not entirely the case in the forthcoming EU General Data Protection Regulation. The definition of personal data has been modified in the draft regulation to include pseudonymised data, where a link to identifiers is maintained but held separately from the dataset.

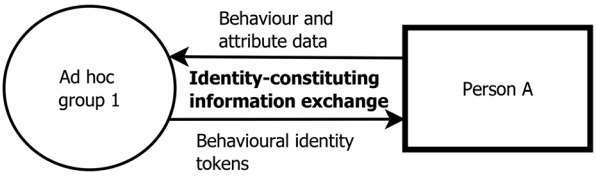

Fig. 1 Information exchanged to create a profiling identity

them (i.e. assigning a higher risk to the group), affect all members of the group including members not yet observed by the analytics system.

Shared ownership of identity is largely ignored in data protection law (cf. Leese 2014; Hildebrandt 2011; Floridi 2012; Taylor et al. 2017). Existing legal protections reflect piecemeal responses to particular egregious uses of shared offline identifiers in decision-making, seen, for instance, in US anti-discrimination law (Barocas 2014) or the ban on personalised insurance premiums based on risk profiling in the EU (Newell and Marabelli 2015). Case-based regulation and detection of discriminatory decision-making is feasible when the set of shared identifiers or groups to be legally protected remains small. The challenge now faced is that analytics routinely creates shared identifiers (i.e. behavioural identity tokens). Patterns and correlations used to group individuals are functionally equivalent to identifiers (e.g. name, address) but are not afforded comparable status under existing data protection law. Proxies for protected attributes are not easy to predict or detect (Romei and Ruggieri 2014; Zarsky 2016), particularly when algorithms access linked datasets (Barocas and Selbst 2015). Profiles constructed from neutral characteristics such as postal code may inadvertently overlap with other profiles related to ethnicity, gender, sexual preference and so on (Macnish 2012; Schermer 2011). Beyond legally protected groups, it remains unclear from the outset the types of behavioural identity tokens and decision-making models that can be produced, and which of these are potentially ethically troubling.[4]

Algorithmic classification may prove ethically problematic for a number of reasons. It can, for example, reduce individual control of identity. The identities of individual members are mediated by knowledge about the group. Groups are formed from observables chosen at a particular level of abstraction for a particular purpose (cf. Floridi 2008). Ad hoc groups are constructed from observables of interest to algorithmic classification systems, which produce a certain kind of knowledge (e.g. a recommendation, rank, classification). These groups are imperfect reflections of the individuals contained within, constrained by the types of question they are designed to answer, and the flaws of the observables (i.e. the data) from which they are constructed. The meaning given to a particular group, and thus imposed on the individual, will not

[4] A particular type of algorithmic classification system—personalisation systems—provides an example as to why algorithmic classification must be considered ethically relevant. Personalisation segments a population so that only some groups are worthy of receiving certain opportunities or information, re-enforcing existing social (dis)advantages. Personalisation systems create self-fulfilling behaviours and limit the opportunities available to users according to their classification within the system (Macnish 2012; Leese 2014). Personalisation "involves unseen, categorical, computational judgments about which searches, articles, or purchases should *probably* come next" (Ananny 2016, 103). Discrimination or social, economic and epistemic benefits of big data can inadvertently localise around groups that offer easy or interesting analysis opportunities (Crawford et al. 2014). Questions of the fairness and the distributive justice of such practices can be raised (Danna and Gandy 2002; Rubel and Jones 2014; Cohen et al. 2014).

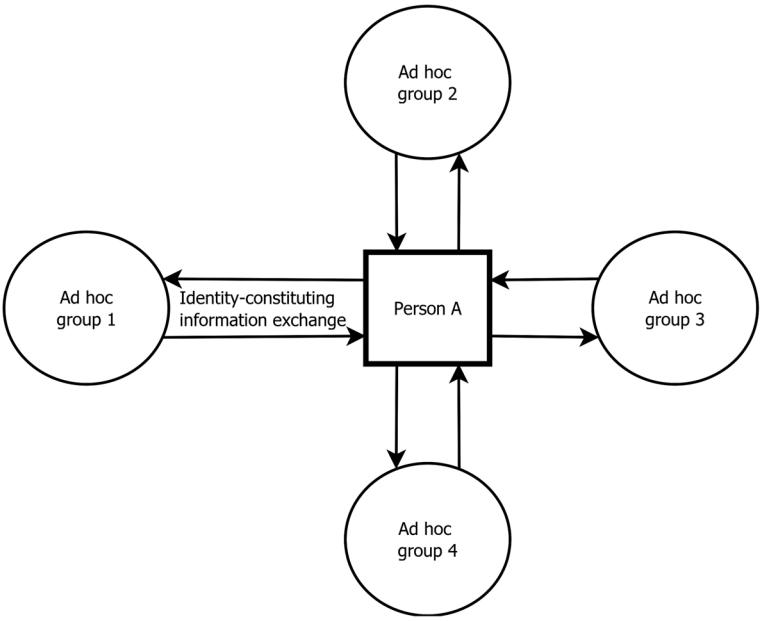

Fig. 2 Profiling identity structure

necessarily reflect her self-understanding (Lupton 2014). Ad hoc grouping provides unanticipated ways of viewing and inferring information about the individual. Algorithmic classification must therefore be considered a threat to data subjects' capacity to shape and control identity.

The fair and equitable treatment of individuals is also undermined by predictive analytics (cf. Newell and Marabelli 2015). If efficiency of a decision-making process driven by analytics is the primary concern, decision-making that proves unfair to

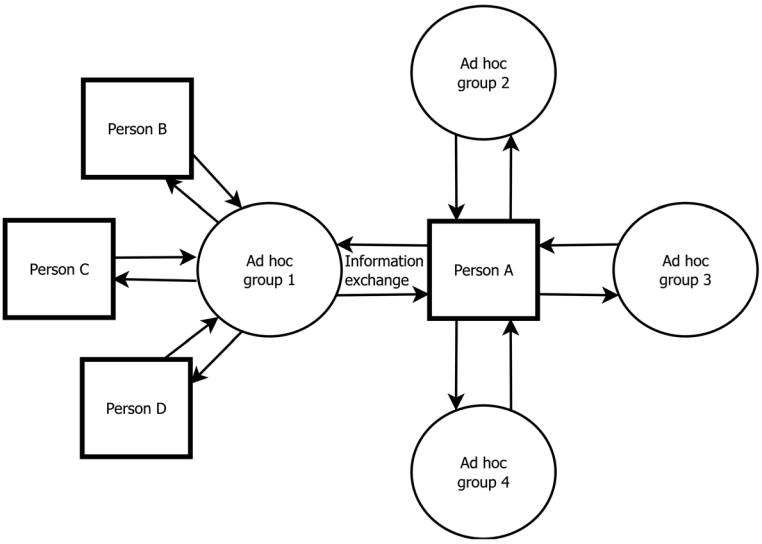

Fig. 3 Shared ownership of behavioural identity tokens

individuals will not necessarily be corrected.[5] Further, it is difficult for lay data subjects to notice and redress this type of unfair or harmful decision-making (Mittelstadt and Floridi 2016). Systematic identification of the limitations of recorded data, or the set of observables that inform algorithmic classification, and how these limitations are reflected in models and decisions produced with analytics (Grindrod 2014) are hindered by the opacity of analytics (Burrell 2016). Ad hoc groups are volatile (Leese 2014); the subject's representation in a classification scheme evolves over time, with new labels applied, tweaked and removed as patterns are identified from new inputs. Memberships can similarly change quickly. Ad hoc groups can overlap with observable groups (e.g. likely voters for a political party), but with membership requirements based on statistical correlations and relevant features of input data rather than attributes apparent to members (e.g. party affiliation). Further, on the basis of irrelevant attributes, individuals may be treated differently from other individuals similar to them in all relevant aspects.

Privacy, conceived as the right to control data about oneself, is an unrealistic ideal under these conditions. The scale of the behavioural data produced in information societies, the complexity of methods to make sense of it (Mittelstadt and Floridi 2016) and the shared ownership of behavioural identity tokens all undermine efforts to protect individual's rights to privacy and identity with limits on processing of identifiable data. Protections are required that respond to the inherent ethical uncertainty of algorithmic classification and the emerging capacities to learn about individuals through knowledge about the groups they are allocated.

3 Group Privacy as the Right to Inviolate Personality

To develop privacy protections responsive to advances in data analytics, an approach to privacy is required that can address groups and shared identity. Group privacy as conceived in the following sections is based on Luciano Floridi's concept of an *informational identity* (Floridi 2011) understood within his broader philosophy of information (Floridi 2013b). When referring to group privacy, I am explicitly addressing the privacy of information that constitutes identity; a violation of informational privacy disrespects an individual or group's claim over information about itself.

This claim can be interpreted in many ways: to control, manage, own or prevent access to information about the self. Traditionally, the damage caused by breaching privacy has been linked in reductionist theories to its undesirable consequences (e.g. distress or discrimination) or in ownership-based theories to the individual's right to exclusive use of information. Problems exist with both approaches (for a discussion, see Floridi 2016a, b).

The alternative approach adopted here to describe group privacy, of privacy as *identity constitutive* (Floridi 2011), connects privacy to the integrity of information

[5] Zarsky (2016, 127–128) explains this point: "As a result of the algorithmic process, individuals might be treated differently than their peers—people similar to them in every *relevant aspect*—on the basis of irrelevant differences…for *some* individuals (yet not enough to render the entire process inefficient), the factors are irrelevant or incorrect. Thus, the process proves to be efficient overall, but in some instances, and for some people, unfair…"

constituting one's identity (Floridi 2016a). While this is certainly not the only feasible theoretical basis for the conceptual development of *group privacy*, it is adopted here due to its suitability to the problem of shared ownership of identity-constituting tokens.

According to this approach, identity is constituted by information describing the individual or group. Privacy is therefore a respect for the right to inviolate personality (Warren and Brandeis 1890, 31) or the "right to immunity from unknown, undesired, or unintentional changes in one's own identity as an informational entity, both actively and passively" (Floridi 2016a). According to this view, identity itself, and the information constituting it, has a value independent of the role it plays in decision-making that produces harmful or beneficial effects. The integrity of the subject's identity is breached when data or information is added to subject's identity without consent. Breaches of identity are considered an attack on the person who is constituted by her information.

A right to inviolate personality protects against the tendency in mature information societies (Floridi 2016b) to produce a longitudinal, semi-permanent view of the individual through data analytics, which would violate her right "to be allowed to experiment with one's own life, to start again, without having records that mummify one's personal identity forever" (Floridi 2006). Compared to offline decision-making, algorithmic classification poses new restrictions on an individual's capacity to shape and alter identity over a lifetime, albeit with historical precedence. Breaches of the integrity of identity are not limited to the online world; rather, the description of a group privacy right in relation to data analytics and automated decision-making is a response to an issue of the scale and persistence of breaches, rather than a fundamentally new type of breach of the integrity of identity. The difference is that analytics allows for the identities or beliefs at the basis of the choice architecture offered to the person not to be limited to a single third party actor, but rather something that persists over time, travels with the person between systems and affects future opportunities and treatment at the hands of others. How a classification is reached is arguably irrelevant to the impact on the data subject's capacity to shape and alter identity over time and the power afforded to decision-making entities to externally shape her identity. The accuracy of a categorisation does not affect its persistence; so long, as the grouping is in place or the categorisation applies to the subject, the subject retains an interest in the information constituting it, even if the grouping proves to be inaccurate.

We can speak of the integrity of identity being breached in two senses. Active breaches of identity involve replication and manipulation of the person's identity within a system. Passive breaches involve indirect pressure placed on the person to include newly generated information in her identity, caused by a failure to obtain meaningful consent for processing (Floridi 2006). Passive breaches describe potential future effects on identity. A group member's self-perception will not necessarily be affected by information about the group due to a lack of self-awareness (i.e. she does not know she is a member) or a small effect size (e.g. the disturbance of a poorly targeted advertisement). The right to inviolate personality is, nevertheless, passively breached in both cases because information that will potentially affect the individual or group's identity in the future has been generated.

Algorithmic classification primarily produces passive breaches of identity. Members of ad hoc groups will struggle to become self-aware regarding the group's existence due to the opacity, secrecy and ubiquity of analytics in information societies (cf. Leese 2014; Hildebrandt 2008). It is for this reason that ad hoc grouping poses a different challenge to

identity management than membership of offline collectives or ascriptive groups (see Section 4.1) in which self-awareness and collective agency are possible. Pragmatically, third party usage of information that constitutes identity is unpredictable; focusing a privacy right on the integrity of identity made up of this information, rather than the effects of its usage, is therefore required to ensure privacy is protected independent of the capacity to observe and correct for harmful effects of data processing.

Floridi's approach is preferred because it formalises the harm to identity caused by third parties that create actionable knowledge about groups. An ad hoc group's identity consists of the classifications and rules constructed by an algorithmic classification system (i.e. why you are a member of this particular group), along with predictions of unobserved or future collective preferences and behaviours. This identity is viewed by the system itself and decision-making processes influenced by its outputs. Opportunities to protect identity require awareness of when, where and how this identity is crafted. Both an individual and group's right to inviolate personality can therefore be violated when identity is crafted externally, without either's consent or awareness.

Floridi compares this process to kidnapping: "the observed is moved to an observer's local space of observation (a space which is remote for the observed), unwillingly and possibly unknowingly. What is abducted is personal information, even though no actual removal of information is in question, but rather only a cloning of the relevant piece of personal information" (Floridi 2006). The production of information that modifies identity is considered a disrespectful attack on identity. Analytics is precisely this: the individual's identity (as contained in the system's inputs) is temporarily kidnapped and modified through classifications that reveal something new about the person (cf. Hildebrandt 2008). External identity construction breaches the individual's inviolate personality.

The privacy right described by Floridi is a fundamental right, meaning the starting point for any negotiation over privacy should start from a position of respecting the right (Floridi 2006). As Floridi (2016a) notes, in contrast to reductionist theories which allow for privacy to be overridden through consequentialist appeals, the identity-constituting approach starts from the position that the right should be respected. Fundamental rights cannot be negated as such by consequentialist appeals but can, nonetheless, be waived through negotiation or consent from the rights-holder or breached due to overriding interests. This is not the same as the reductionist approach: where consent is not obtained, meaning "unknown, undesired, or unintentional" changes to identity occur, the subject's privacy right is breached. Grounds for justifiable breaches under special circumstances can of course be identified, but this does not alter the fact that a breach of privacy has occurred.

It is questionable whether agreement mechanisms governing many digital sources of data (e.g. privacy policies and *terms and conditions*) grant meaningful consent. Similarly, consent does nothing to address the cumulative effects of knowledge generated through analytics on social and organisational decision-making processes (e.g. healthcare commissioning). Processing of aggregated or anonymised data remains unfettered.

4 Group Privacy in Practice

Group privacy as a right to inviolate personality aims to protect the integrity of a group's identity, which is non-reducible (i.e. not merely a sum of the identities of

members). If protections of group privacy are introduced as a way to correct the imbalance of power created by the proliferation of data analytics (i.e. the power of third parties to shape and control a data subject's identity), three interpretations of the concept come to mind.

A *strong interpretation* would allow for rights-holders to make claims over any data processing that generates identity-constitutive information via algorithmic classification (cf. Altman 1977). A strong right would equate to the control of third party data processors and algorithmic classification systems. This approach would create new conflict points for data processors and subjects.

A weakened but practically feasible interpretation recognises the claims of rights-holders over identity-constitutive data processing as valid but limits them to oversight, not control. A duty would thus be created for data processors to keep rights-holders *in the loop*. This *moderate interpretation* differs from existing rights and duties enshrined in EU data protection law in the sense that it validates the individual's claim to information about the groups to which she nominally belongs, either as a rights-holder or member of a rights-holding group.

A *weak interpretation* denies the rights-holders any claim to control how third parties form views about her. On this view, a weak right to group privacy would be one of enhancement only, meaning data processors have a non-binding duty to educate data subjects about how their *data trails* mediate their experiences within analytics-driven decision-making.

4.1 Types of Rights and Rights-Holders

Note that the term 'rights-holder' is preferred here to 'group' or 'individual' in recognition that a right to group privacy can feasibly be held by either. To establish ad hoc groups as legitimate holders of a right to group privacy, it is first necessary to examine in what sense groups can have interests and rights. We can speak of two types of rights: moral and legal. Moral rights establish duties for agents to respect the interests of a patient. In turn, legal rights are moral rights enacted by law, specifying protected interests of moral patients and concomitant duties for moral agents (Jones 2016). Here, we will first consider whether ad hoc groups can be granted a moral right to privacy. Later, we will examine how a moral right can be enacted as a legal right to group privacy (see Section 5).

The nature and validity of group rights is a long-standing problem in political philosophy (Jones 2016). The debate does not concern whether certain groups should have rights and interests in practice, but rather which types of groups are valid rights-holders. Rights held by groups (i.e. a *group right*) are not reducible to the rights of members; it is a right held by a group qua group. Members of the group can also be granted individual rights due to their membership in the group, seen, for instance, in rights and privileges granted to members of a social club. Such rights are not group rights as discussed here; rather, a group right is attributed to the group itself, not individual members (Floridi 2016a).

According to how membership is established, three types of groups can be recognised as potential rights-holders:

- *Collectives*—a group intentionally joined due to collective interests, shared background or other explicit common traits and purposes (examples *patient advocacy group*, *labour unions*)

- *Ascriptive groups*—a group whose membership is determined by inherited or incidentally developed characteristic. Ascriptive groups cannot normally be intentionally joined or left without redefining the boundaries of the group (examples *ethnic groups*, *patient cohorts*)
- *Ad hoc groups*—a group whose membership is assembled for a third party interest according to perceived links between members, often for a time- or purpose-limited period with volatile membership requirements (examples *market segments*, *profiling groups*)

Collectives and ascriptive groups are already legally recognised as legitimate rights-holders in some contexts. Collective rights include the right to self-determination held by nations, or legal rights granted to corporations (List and Pettit 2011). Prior work on group privacy has established a right for labour unions, understood as collectives, to assemble (Bloustein 1976). Anti-discrimination law protects groups or *protected classes* from *disparate impact* in social and political decision-making (Barocas and Selbst 2015). Other examples include prohibitions on genetic discrimination (Natowicz et al. 1992) and protections of minority cultural and linguistic traditions. In contrast, ad hoc groups are not yet recognised as deserving of legally enshrined privacy rights.

Rights can be granted to groups on the basis of several characteristics. As revealed by the aforementioned examples of rights held by collectives and ascriptive groups, collective identity and agency are often considered minimal requirements for the possession of group rights.

Collective identity can refer to both a reducible identity aggregated from the independent identities of members (such that the group's identity necessarily changes when members leave) and a non-reducible identity held by the group itself. Institutions show how collectives can possess this latter type of non-reducible group identity; sports clubs, for instance, have an identity (e.g. history, norms) and interests independent of the identities and interests of their membership which changes over time (French 1984).

Collective agency refers to the capacity to act to fulfil or protect the interests of the group. Needs and interests are shared between members insofar as "what benefits or harms any member must benefit or harm all" (Jones 2016). Agency requires self-awareness: groups cannot act in their best interests without first being aware that the group exists. All members of a collective will by definition possess self-awareness. However, for ascriptive groups, self-awareness is required only at the group level; genealogy, for instance, demonstrates how members of an ascriptive group can be ignorant of inherited traits and cultural history.

4.2 Ad Hoc Groups as Rights-Holders

Ad hoc groups lack both collective identity and agency. They are assembled by algorithmic classification systems as *types* of individuals with common *characteristics* (Floridi 2014) as opposed to common *intentions*, which suggests a lack of self-awareness, collective interests or agency. Ad hoc groups are created by a third party that links a set of individuals according to perceived similarities (French 1984). An identity can be imposed on members unaware that they belong to the group. The third party interests in the formation of ad hoc groups also need not be long term: ad hoc groups need last only long enough to answer a question posed by a data processor. Both

characteristics undermine any ascription of collective agency or intentionality.[6] Ad hoc groups are thus unique insofar as they are externally crafted and lack the intentionality and self-awareness of legally protected groups.

Collective identity and collective agency are not, however, absolute requirements for groups to hold rights. Instead, moral patienthood may prove sufficient. For a group itself to hold a right, it must "possess a moral standing that is not reducible to the standing of its members" to possess rights (Jones 2016). Moral patienthood establishes precisely this moral standing.

Thresholds for patienthood vary. A strict anthropocentric view grants patienthood only to fellow moral agents (cf. Wellman 1995). In contrast, the mere possession of a "reason to regard a group as an object of moral concern" (Jones 2016), such as the recognition of collective interests, may itself provide sufficient reason to treat an entity as a moral patient. Similarly, particular capacities can serve as a minimal threshold, for instance the capacity to suffer (cf. Singer 1993) or be harmed by entropy as an informational object (cf. Floridi 2013a).

On all but the strict anthropocentric view, ad hoc groups can be moral patients (Floridi 2013b). Floridi (2016a) argues that groups can be rights-holders by the same logic that rights are attributed to individuals, so long as the group itself is treated as an individual (cf. corporate personhood). According to him, groups serve the same role as individuals in interactions involving the exchange of information. Rights can simultaneously be held by the individual and the individual's group as long as identity-forming interactions exist at both levels. If these interactions constitute the group's identity (as is the case for individuals), then groups can be recognised as possessing an interest in controlling that identity. This interest can be formally recognised as a group right to inviolate identity (Floridi 2016a), even if the group cannot itself act to control that identity. Agency can be hypothetically attributed to ad hoc groups if third party processors act to respect interests the group would presumably have if self-aware. Presumed interests can be based upon observed interests of individual members, for instance the interest to be accurately classified according to their data. This is not to say the group's interests are reducible to the interests of members, but rather that an educated guess can be made about what the group would value if given the opportunity to assemble and act collectively.

In the case of ad hoc groups, identity-forming interactions involving the group and individual members intersect due to shared ownership of behavioural identity tokens (see Figs. 1, 2 and 3). The behaviours of individuals produce data used to define an actionable group. Actions taken by members affect the interests of the whole group due to the predictive power attributed to the algorithmic classifications; the group is actionable in the sense that members have been observed (or perceived) to be similar and are thus predicted to behave or respond to stimuli (e.g. advertisements) similarly in the future. As a result, the group as a whole has a claim to the tokens used to define it and guide future actions taken towards it. Individual members do not retain an exclusive claim to their data trails once algorithmically grouped.

[6] With that being said, persistent profiling identities can be extremely valuable in analytics. The profiling identity, as a record of one's ad hoc group memberships, is akin to a digital *permanent record* that reflects one's identity within one or more linked algorithmic classification systems. An electronic health record linked across GPs, hospitals and other facilities that describes a patient's personal health risks is but one example.

From this brief discussion, ad hoc groups can tentatively be treated as moral patients and thus valid rights-holders.[7] Regardless of practical feasibility, this conclusion helpfully emphasises the co-existence of non-reducible group and individual identities, as well as privacy interests. However, even if granted rights, ad hoc groups cannot act to protect their interests due to a lack of collective agency and self-awareness. Consideration is therefore required of how to enact formal protections of group privacy on behalf of ad hoc groups.

5 Two Implementation Models

Two existing legal instruments demonstrate how the interests of groups can be protected when collective action is infeasible: American class action and anti-discrimination law. In class action lawsuits, a group (represented by a plaintiff) seeks compensation from an organisation based around a violation of the rights or interests of the group (e.g. an interest in product safety). Similar to algorithmic classification, the group does not exist prior to the lawsuit; the lawsuit provides a *purpose* around which the group is assembled.

Let us consider the hypothetical example of a fault identified in the safety harnesses of a particular model of automobile. In this case, the group consists of all owners of that type of car. The class action group does not exist before the lawsuit in the sense that a reason does not yet exist for it to be formed. The formation of class action groups, as with ad hoc groups, is abstract; the group does not meet or interact but is rather legally assembled by a third party (the plaintiff) for a well-defined purpose (the lawsuit). The plaintiff seeks compensation on behalf of the group; similarly, the data processor undertakes knowledge work on behalf of the group that has been assembled by a classifier algorithm. Critically, only the group can act to pursue its interests; individual members cannot separately pursue personal interests within a class action lawsuit.

A similar mechanism can be found in American anti-discrimination law concerning disparate impact. According to the principle of disparate impact,[8] actions in employment, housing, policing and other areas are considered discriminatory and illegal if protected classes of people (e.g. groups based on race, gender, religion) suffer disproportionate adverse impact compared to non-members.[9] Critically, the legality of the action is determined by the distribution of impact across a population (i.e. not only the

[7] This issue requires further contextualisation within theories of political and legal rights, which goes beyond the scope of this paper. Furthermore, it may largely be a semantic point due to ad hoc groups lacking self-awareness and collective agency. In both cases, the possession of a moral right creates duties for agents to respect the rights-holders' interests in inviolate personality. Even if a right to group privacy is only held by individual members, the possession of equivalent rights by all group members provides a check on the power of any individual member to exercise control over the group's identity.

[8] An in-depth discussion of disparate impact and detection of discrimination in data analytics can be found in Barocas and Selbst (2015). A discussion of the legal origin of the principle of disparate impact in the Civil Rights Act of 1964 is provided by Rutherglen (1987).

[9] The discussion of anti-discrimination law is limited here to the principle of disparate impact. However, as Zarsky (2016) notes, at least two other types of discrimination can arise in analytics: (1) the explicit usage of protected characteristics (e.g. race) to inform decisions and (2) the usage of biased or skewed datasets that discriminate against protected groups, often resulting from "human biases in measurement or other past wrongs that might lead to over-representation of some forms of negative data about minorities". For a discussion of further types of discrimination in analytics, see Romei and Ruggieri (2014).

affected groups). The method and rationale of the action are irrelevant to its legality. Attention is paid only to the effects of the action to prevent discrimination against protected classes through proxy features (e.g. post code as a proxy for ethnicity).[10]

Disparate impact doctrine differs from algorithmic classification in that protected classes are legally well-defined prior to the identification of adverse impact. While the characteristics driving the search for adverse impact are defined in advance, the actual groups are not; the group 'African Americans' is a legally protected class, but a group relevant to disparate impact analysis, such as 'African Americans subjected to disproportionate adverse impact in social housing in Brooklyn', is not assembled before adverse impact is identified.

Both examples provide precedents for the protection of group privacy interests. Group privacy is closely related to protection of equality through anti-discrimination law insofar individuals are protected in both cases from unfair treatment on the basis of actual or inferred attributes. However, group privacy based on the right to inviolate personality specifically aims to protect against (1) unwarranted third party manipulation of identity and (2) harms from automated decision-making based upon profiling identities assembled by third parties.

Ad hoc groups differ from protected classes in that they need not be based on intuitive features or parameters (Burrell 2016). Analytics allows for a new type of group to be formed that is not protected by anti-discrimination provisions, because the groups need not align with existing protected classes or attributes, such as those addressed in fairness- and discrimination-aware analytics (e.g. Romei and Ruggieri 2014; Kroll et al. 2016; Ajunwa et al. 2016). Rather, the defining features of an ad hoc group are algorithmically determined and, thus, not necessarily observable or interpretable to humans. Anti-discrimination mechanisms built around offline identifiers are thus insufficient to protect groups constructed by algorithmic systems from harmful decisions based on attributes that are not merely proxies for legally protected attributes (e.g. ethnicity, gender). The future development of priacy-, fairness-, and discrimination-aware analytics techniques should address this fact and expand their scope of concern beyond legally protected classes and proxies thereof.

Generically, protections of group privacy can assist in predicting, detecting and remedying invasive classifications and in preventing access to some data types for particular purposes. Class action and disparate impact point towards a possible implementation model for group privacy protections, i.e. *reactive* recourse mechanisms to detect and compensate for violations of a group's privacy interests.

Reactive protections are concerned with ex post detection of classifications that attack a group's informational identity. Reactive protections can (1) provide recourse or compensation to groups when invasive classifications are detected or (2) correct classification systems that subject groups to disproportionate adverse impact (e.g. targeted risk profiling). As with individual informational privacy, oversight allows group members to anticipate and respond to the ways in which beliefs about the group mediate their experiences as individuals.[11] Oversight in this sense would empower

[10] Targeted policing through an algorithmic analysis of crime statistics provides an example; focusing on crime *hot spots* can disproportionately expose particular ethnic and socioeconomic groups to police scrutiny (cf. Harcourt 2006). The rationality and desirability of preventing crime by targeting high crime areas have no bearing on whether such systems are considered discriminatory.

[11] These interests are supported by requirements in the EU Data Protection Directive (95/46/EC) and forthcoming General Data Protection Regulation for data processors to explain the decision-making logic of automated processing when queried (European Commission 2012).

groups to assess the acceptability of algorithmic classification for different processing contexts and data types.

Proactive protections are also feasible to (1) predict and prevent the assembly of certain types of profiles or groups, (2) prevent classifier algorithms from being applied to certain types of questions or purposes or (3) prevent classifier algorithms from being provided with inputs containing protected attributes (Barocas and Selbst 2015), such as gender or ethnicity (Calders et al. 2009; Kamiran and Calders 2010; Schermer 2011). Such protections address the problem of information overload, the overwhelming effort required to control one's data in the information age, by setting particular data types and questions off limits.

5.1 Challenges

Both models face several challenges. Proactive protections may be preferable due to ad hoc groups lacking self-awareness and collective agency. Recourse mechanisms cannot be accessed by individuals unaware they are members. A collective voice does not exist within ad hoc groups. This challenge is overcome in anti-discrimination law, which proactively restricts processing of protected types of data and processing for discriminatory purposes. Demographic groups cannot assemble and debate the merits of processing; control is exercised as a legislative level (often including representatives of the group that can speak, presumably coherently, about the interests of the group), not by individual members of the group.

While nominally a vote in favour of proactive protections, this lack of unified voice can also support a reactive model. Individual and group interests in privacy can conflict. Groups can make judgments that are not derived or equivalent to the sum of members' individual judgments (List and Pettit 2011). However, deliberation cannot occur within a group whose members are not aware of one another. As a result, prohibitions can easily come to unfairly represent a particularly motivated individual's views rather than the interests of the group. Algorithms and ex post accountability mechanisms can be designed to protect the privacy of individuals, without offering any protection to groups or group identity (i.e. the group's inviolate personality). Prohibition of certain types of classification or inputs is also risky when their potential utility is uncertain; beneficial processing could easily be blocked by blunt regulation (cf. Nissenbaum 2004).

Anti-discrimination law demonstrates the difficulty of blanket prohibitions on analytics to protect group privacy interests. The EU takes a principled systematic approach to anti-discrimination, identifying types of discrimination as illegal across all data processing contexts. Elsewhere, for instance in the USA or Australia, a piecemeal approach is taken, with specific forms of discrimination outlawed in response to problems arising in specific contexts, such as housing or credit decisions (Romei and Ruggieri 2014). Rational preventative measures to protect group privacy would take a principled, universal approach, given the non-territorial, trans-national and persistent nature of data analytics and the models and profiles produced therein.

One response to the difficulties faced by proactive protections is to limit group privacy protections to recourse; rather than attempt to predict something that is highly uncertain, effort should focus on developing methods to identify and compensate violations of a group's privacy interests. Class action and disparate impact both provide

models for group recourse. As a reactive concept, group privacy provides an alternative route for individuals to maintain their informational identity, even when the link between the individual's privacy and the data has been lost through anonymisation.

Reactive protections are not without comparable challenges. A primary problem with reactive group privacy concerns the difficulty of detecting violations of group interests. Detection is made possible in anti-discrimination and class action law because the areas to which they are applicable (e.g. social housing, product safety) and the groups of people they are intended to protect (i.e. legally protected classes) are well defined; the search for disparate impact can start, for instance, with housing statistics for protected groups in a particular geographical location.

Further work is required to establish initial anchor points for the detection of group privacy violations.[12] Generically, a violation occurs when the group's identity is changed by unwanted information. How this occurs in practice is not well established. Prior work on the rights of collectives in international law provides some idea (e.g. Bloustein 1976; Bisaz 2012). High profile examples of invasive profiling are also helpful. However, these sources alone are insufficient to develop a general taxonomy of violations. Ethically problematic profiling tends to only be identified when the system fails spectacularly and in public. Mundane violations that indicate a cumulative effect on identity (e.g. displaying the wrong advertisement) are harder to record. Auditing functions built into analytics systems may go some way to solving this problem by allowing for ex post detection of violations through examination of the inputs and decision-making rules resulting in a problematic decision (cf. Sandvig et al. 2014; Zarsky 2016; Mittelstadt et al. 2016).

Detection mechanisms, however, face problems stemming from the complexity and opacity of analytics systems (Zarsky 2016). As discussed above (see Section 4.2), ad hoc groups are comparable to the type of groups and interests protected in class action and anti-discrimination law. However, they differ in one critical aspect. Disparate impact occurs when (1) evidence of disproportionate adverse impact is uncovered, affecting (2) a well-defined group (e.g. a protected class). Uncertainty is high for the first component (evidence can be sought in innumerable places) but low for the second because the law applies only to particular protected groups. Disparate impact is a feasible recourse mechanism precisely because the affected groups are defined according to intuitive characteristics (e.g. ethnicity, product ownership). Ad hoc groups are not similarly intuitive, defined instead by machine interpretable correlations (cf. Burrell 2016; Lisboa 2013). A comparable form of recourse for group privacy violations would be substantially more complex because both components are highly uncertain, owing to the opacity of analytics systems.

Constructing appropriate channels for feedback and recourse is, nonetheless, particularly important to protect group privacy in the future. Understanding how automated and bureaucratic decision-making amplifies the political power of decision-making organisations and disempowers data subjects in the process will be a key to designing appropriate recourse mechanisms. The complexity of automated decision-making amplifies the impact of classifications on both individual and group identity. Interpretable or low-dimensional decision-making can similarly restrict an individual's identity

[12] For a discussion of the lack of evidence of harms arising from data misuse in the context of biomedical processing, see Laurie et al. (2014).

over time if a classification is persistent; being classified as 'black' on the basis of one's ancestry can, for instance, have equivalent effects on future choices and treatment by others, without involving complex decision-making methods. However, highly complex and poorly interpretable algorithmic decision-making adds an extra barrier to controlling one's identity in the face of such classifications. The subject must overcome not only the bureaucratic and organisational power that makes such classifications persistent over time, but is also less able to challenge the classification if its rationale or underlying evidence cannot be understood and, therefore, questioned. The forthcoming General Data Protection Regulation addresses this problem insofar as it specifies vague rights and duties for data subjects and data controllers concerning automated decision-making. The eventual interpretation of these provisions will ideally build upon an understanding of the relationship between automated and bureaucratic decision-making.[13]

One way to address these challenges is to identify processing contexts, purposes and data types that allow violations of inviolate personality to be negotiated with affected group(s). The relative strength of a group's privacy interests will vary between processing contexts. While the right to inviolate personality is absolute in principle, in practice, it must be balanced with individual privacy rights and the social, commercial and epistemic benefits of data processing. It is thus not, in principle, an additional barrier to data processing but rather a third set of considerations for ethical assessment of analytics platforms. Group privacy could, for instance, support the introduction of consultations with patient cohorts as a complement to existing individualistic privacy and consent mechanisms in biomedical research involving analytics.

6 Conclusion

Advances in data analytics necessitate new protections for the privacy interests of ad hoc groups formed by algorithmic classification. Mechanisms are required to protect the privacy interests of groups independent of the interests of their individual members. The privacy interests of ad hoc groups are not reducible to the interests of members due to shared ownership of behavioural identity tokens. It is these tokens, as opposed to information about a unique individual, that mediate the individual's treatment within a given analytics system. Each individual thus has a valid claim to the tokens constituting the group's identity.

The group's identity is not, however, reducible to the identities of individual members, which are defined by a far greater set of identity tokens than those that define the group. If an ad hoc group possesses a non-reducible identity, interests can be ascribed to the maintenance of this identity. Protecting privacy by hiding the identity of the subject of processing does not therefore address how algorithmic classifiers make sense individuals.

As suggested by the depth of philosophical work on group rights,[14] establishing ad hoc groups as valid rights-holders is a too complex problem to fully explore here. The intention here has not been to make this case in full, but rather to argue that ad hoc

[13] This is a complex but important issue that goes beyond the scope of this paper.
[14] See Jones (2016) for an overview.

groups possess privacy interests not reducible to those of members, despite lacking collective agency and self-awareness. If it is decided that ad hoc groups cannot hold rights, the protections described here (see Section 5) would need to be linked to or based on the rights of individuals, or an ethical duty imposed on data processors. This important theoretical issue requires further attention.

The forthcoming EU General Data Protection Regulation includes provisions for data subjects to request information about the logic involved in automated, algorithmic decision-making. This may prove a turning point for legal recognition of group privacy rights.[15] As shown by the difficulty of detecting group privacy violations, practical implementation of this right will be extremely difficult. Going forward, ad hoc groups should be formally recognised as moral patients in data protection law and privacy theory. The privacy interests of ad hoc groups must be recognised to ensure privacy adapts to advances in the ways we make sense of each other through the data we create.

Compliance with ethical standards

Funding This study was funded by the PETRAS IoT Hub - a EPSRC project (grant no. EP/N023013/1).

References

Ajunwa, I., Friedler, S., Scheidegger, C. E., & Venkatasubramanian, S. (2016). *Hiring by algorithm: predicting and preventing disparate impact. SSRN Scholarly Paper ID 2746078*. Rochester: Social Science Research Network http://papers.ssrn.com/abstract=2746078.

Altman, I. (1977). Privacy regulation: culturally universal or culturally specific? *Journal of Social Issues, 33*(3), 66–84.

Ananny, M. (2016). Toward an ethics of algorithms convening, observation, probability, and timeliness. *Science, Technology & Human Values, 41*(1), 93–117. doi:10.1177/0162243915606523.

Barocas, Solon. 2014. Data mining and the discourse on discrimination. https://dataethics.github.io/proceedings/DataMiningandtheDiscourseOnDiscrimination.pdf.

Barocas, Solon, and Andrew D. Selbst. 2015. Big data's disparate impact. SSRN Scholarly Paper ID 2477899. Rochester: Social Science Research Network. http://papers.ssrn.com/abstract=2477899.

Bisaz, Corsin. 2012. *The concept of group rights in international law: groups as contested right-holders, subjects and legal persons*. Martinus Nijhoff.

Bloustein, E. J. (1976). Group privacy: the right to huddle. *Rutgers Camden Law Journal, 8*, 219.

Burrell, J. (2016). How the machine "thinks:" understanding opacity in machine learning algorithms. *Big Data & Society.* doi:10.1177/2053951715622512.

Calders, Toon, Faisal Kamiran, and Mykola Pechenizkiy. 2009. Building classifiers with independency constraints. In *Data Mining Workshops, 2009. ICDMW'09. IEEE International Conference on*, 13–18. IEEE.

[15] It is worth noting that this right has existed in some form the Data Protection Directive 95/46/EC. To the author's knowledge, no member state has yet implemented a workable mechanism for individuals to query automated decision-making.

Cohen, I. G., Amarasingham, R., Shah, A., Xie, B., & Lo, B. (2014). The legal and ethical concerns that arise from using complex predictive analytics in health care. *Health Affairs, 33*(7), 1139–1147. doi:10.1377/hlthaff.2014.0048.

Crawford, K., Gray, M. L., & Miltner, K. (2014). Critiquing big data: politics, ethics, epistemology | special section introduction. *International Journal of Communication, 8*, 10.

Danna, A., & Gandy Jr., O. H. (2002). All that glitters is not gold: digging beneath the surface of data mining. *Journal of Business Ethics, 40*(4), 373–386.

European Commission. 2012. Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation). COM(2012) 11 final. Brussels: European Commission. http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf.

Floridi, L. (2006). Four challenges for a theory of informational privacy. *Ethics and Information Technology, 8*(3), 109–119.

Floridi, L. (2008). The method of levels of abstraction. *Minds and Machines, 18*(3), 303–329. doi:10.1007/s11023-008-9113-7.

Floridi, L. (2011). The informational nature of personal identity. *Minds and Machines, 21*(4), 549–566. doi:10.1007/s11023-011-9259-6.

Floridi, L. (2012). Big data and their epistemological challenge. *Philosophy & Technology, 25*(4), 435–437. doi:10.1007/s13347-012-0093-4.

Floridi, L. (2013a). *The ethics of information*. Oxford: OUP Oxford.

Floridi, Luciano. 2013b. The philosophy of information. Reprint edition. OUP Oxford. Oxford.

Floridi, L. (2014). Open data, data protection, and group privacy. *Philosophy & Technology, 27*(1), 1–3. doi:10.1007/s13347-014-0157-8.

Floridi, Luciano. 2016a. Three problems with group privacy and their solutions. In Group privacy: new challenges of data technologies. Vol. forthcoming. Philosophical Studies. New York: Springer.

Floridi, L. (2016b). Mature information societies—a matter of expectations. *Philosophy & Technology, 29*(1), 1–4. doi:10.1007/s13347-016-0214-6.

French, P. A. (1984). *Corporate and collective responsibility*. New York: Columbia University Press.

Grindrod, P. (2014). *Mathematical underpinnings of analytics: theory and applications*. Oxford: OUP Oxford.

Harcourt, Bernard E.. 2006. *Against prediction: profiling, policing, and punishing in an actuarial age*. University of Chicago Press. http://bibliovault.org/BV.landing.epl?ISBN=9780226316147.

Hildebrandt, Mireille. 2008. Defining profiling: a new type of knowledge? In *Profiling the European citizen*, edited by Mireille Hildebrandt and Serge Gutwirth, 17–45. Springer Netherlands. http://link.springer.com/chapter/10.1007/978-1-4020-6914-7_2.

Hildebrandt, M. (2011). Who needs stories if you can get the data? ISPs in the era of big number crunching. *Philosophy & Technology, 24*(4), 371–390. doi:10.1007/s13347-011-0041-8.

Jones, Peter. 2016. Group rights. In *The Stanford encyclopedia of philosophy*, Summer 2016 Edition (forthcoming). http://plato.stanford.edu/archives/sum2016/entries/rights-group/.

Kamiran, Faisal, and Toon Calders. 2010. Classification with no discrimination by preferential sampling. In *Proc. 19th Machine Learning Conf. Belgium and The Netherlands*. http://wwwis.win.tue.nl/~tcalders/pubs/benelearn2010.

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Harlan, Y. (2016). *Accountable algorithms. SSRN Scholarly Paper ID 2765268*. Rochester: Social Science Research Network http://papers.ssrn.com/abstract=2765268.

Laurie, Graeme, Kerina H. Jones, Leslie Stevens, and Christine Dobbs. 2014. A review of evidence relating to harm resulting from uses of health and biomedical data. Nuffield Council on Bioethics. http://nuffieldbioethics.org/wp-content/uploads/FINAL-Report-on-Harms-Arising-from-Use-of-Health-and-Biomedical-Data-30-JUNE-2014.pdf.

Leese, M. (2014). The new profiling: algorithms, black boxes, and the failure of anti-discriminatory safeguards in the European Union. *Security Dialogue, 45*(5), 494–511. doi:10.1177/0967010614544204.

Lisboa, Paulo JG. 2013. Interpretability in machine learning—principles and practice. In *Fuzzy logic and applications*, 15–21. Springer. http://link.springer.com/chapter/10.1007/978-3-319-03200-9_2.

List, C., & Pettit, P. (2011). *Group agency: the possibility, design, and status of corporate agents*. Oxford: Oxford University Press.

Lupton, D. (2014). The commodification of patient opinion: the digital patient experience economy in the age of big data. *Sociology of Health & Illness, 36*(6), 856–869. doi:10.1111/1467-9566.12109.

Macnish, K. (2012). Unblinking eyes: the ethics of automating surveillance. *Ethics and Information Technology, 14*(2), 151–167. doi:10.1007/s10676-012-9291-0.

Mittelstadt, B. D., & Floridi, L. (2016). The ethics of big data: current and foreseeable issues in biomedical contexts. *Science and Engineering Ethics, 22*(2), 303–341. doi:10.1007/s11948-015-9652-2.

Mittelstadt, Brent, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. The ethics of algorithms: mapping the debate. *Big Data & Society*.

Natowicz, M. R., Alper, J. K., & Alper, J. S. (1992). Genetic discrimination and the law. *American Journal of Human Genetics, 50*(3), 465–475.

Newell, S., & Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: a call for action on the long-term societal effects of "datification". *The Journal of Strategic Information Systems, 24*(1), 3–14. doi:10.1016/j.jsis.2015.02.001.

Nissenbaum, H. (2004). *Privacy as contextual integrity. SSRN scholarly paper ID 534622*. Rochester: Social Science Research Network http://papers.ssrn.com/abstract=534622.

Romei, A., & Ruggieri, S. (2014). A multidisciplinary survey on discrimination analysis. *The Knowledge Engineering Review, 29*(05), 582–638. doi:10.1017/S0269888913000039.

Rubel, A., & Jones, K. M. L. (2014). *Student privacy in learning analytics: an information ethics perspective. SSRN scholarly paper ID 2533704*. Rochester: Social Science Research Network http://papers.ssrn.com/abstract=2533704.

Rutherglen, G. (1987). Disparate impact under title VII: an objective theory of discrimination. *Virginia Law Review, 73*(7), 1297–1345. doi:10.2307/1072940.

Sandvig, Christian, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2014. Auditing algorithms: research methods for detecting discrimination on internet platforms. *Data and discrimination: converting critical concerns into productive inquiry*. http://social.cs.uiuc.edu/papers/pdfs/ICA2014-Sandvig.pdf.

Schermer, B. W. (2011). The limits of privacy in automated profiling and data mining. *Computer Law & Security Review, 27*(1), 45–52. doi:10.1016/j.clsr.2010.11.009.

Singer, P. (1993). *Practical ethics*. Cambridge: Cambridge University Press.

Taylor, Linnet, Luciano Floridi, and Bart van der Sloot, eds. 2017. Group privacy: new challenges of data technologies. 1st ed. Philosophical Studies. New York: Springer.

van der Sloot, B. (2014). Privacy in the Post-NSA Era: Time for a Fundamental Revision? http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2432104.

Vedder, A. (1999). KDD: the challenge to individualism. *Ethics and Information Technology, 1*(4), 275–281.

Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard Law Review, 4*(5), 193–220. doi:10.2307/1321160.

Wellman, C. (1995). *Real rights*. New York: Oxford University Press.

Zarsky, Tal. 2013. Transparent predictions. University of Illinois Law Review 2013 (4). http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2324240.

Zarsky, T. (2016). The trouble with algorithmic decisions an analytic road map to examine efficiency and fairness in automated and opaque decision making. *Science, Technology & Human Values, 41*(1), 118–132. doi:10.1177/0162243915605575.