4

Big Data as Tracking Technology and Problems of the Group and Its Members

Haleh Asgarinia

Use of Big Data Analytics to Target Persons or Groups

Digital data help data scientists and epidemiologists track and predict outbreaks of disease. Mobile phone GPS data, social media data, or other forms of information updates as epidemics progress are used by epidemiologists to recognize disease spread among specific groups of people. Given the gravity of the risk that certain groups are exposed to, restriction of movement or surveillance could be imposed on them, as we have seen in recent years. In order to control outbreaks of disease, quarantine decisions are taken based on tracking the transmission of the disease on the group level (Taylor 2016). For example, new data sources have been employed in high-stakes scenarios to track a range of life-threatening diseases, including cholera in the wake of the 2010 Haiti earthquake (Bengtsson et al. 2011), malaria transmission via network analysis (Tatem et al. 2009), and Covid-19 (this volume).

In the case of the 2010 cholera epidemics, anonymous cell phone data were used to track and predict cholera epidemics in Haiti after the 12 January 2010 earthquake. Researchers used call records to investigate population movements after cholera struck coastal towns and surrounding areas, demonstrating that many who left these areas moved to cities. This knowledge was crucial because people leaving cholera-affected areas carried the disease with them (Bengtsson et al. 2011). Mobile phone records have also provided a valuable data source for characterizing malaria transmission, enabling policymakers to modify and implement strategies for further preventing transmission (Liu et al. 2012). Using data from mobile phone networks to track population movements has therefore helped improve responses to disasters and disease outbreaks.

More recently, in response to the Covid-19 crisis, big data analytics helped public officials in making decisions about how to reopen society safely and how much activity to allow. To accomplish this, epidemiological models that capture the effects of changes in mobility on virus spread have been developed by

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

Haleh Asgarinia, *Big Data as Tracking Technology and Problems of the Group and Its Members* In: *The Ethics of Surveillance in Times of Emergency.* Edited by: Kevin Macnish and Adam Henschke, Oxford University Press.
© Oxford University Press 2023. DOI: 10.1093/oso/9780192864918.003.0005

reflecting on patterns of human interaction at non-residential locations of interest, such as shops, restaurants, and places of worship. The results of such findings could be used to infer which activities should be continued and which should be avoided. According to the model, infections in venues such as restaurants, gyms, and religious establishments play a disproportionately large role in driving up infection rates, restricting the reopening of such establishments, and making them a key target for control (Chang et al. 2021). As a result, big data analytics as tracking technologies can help authorities control and manage the Covid-19 pandemic and bring a premature end to epidemics.

Policymakers and authorities use information derived from big data analytics to target groups or persons. When an entity is the target of information, this means that observers, policymakers, or authorities have information that they relate to an entity in the world (Henschke 2017). The observer uses the information to target the person who is infected and the person who is at risk of infection because they have been in contact with the infected person. Moreover, the observer can target groups as potential carriers of a disease, rather than addressing persons as patients.

Though promising, pandemic surveillance brings a series of challenges for those targeted by the information derived from big data analytics. Targeting persons and groups risks causing harm to a person, as a member of a group, and to a group *qua* group. Three of the ethical issues raised by targeting a person with the information generated at the aggregate level are consent, social justice and fairness, and privacy. The negative consequences of data processing at group level are the risks of group discrimination or stigmatization. In these types of cases, the problem is not that this or that specific person has been harmed, but that the group as a whole is affected and thereby undermined (Sloot 2017). These ethical issues and harms will be discussed in the following sections.

The EU General Data Protection Regulation (GDPR) is considered a key to the successful development of technologies to tackle the Covid-19 pandemic (Mikkelsen, Soller, and Strandell-Jansson 2020; Newlands et al. 2020). The consent of a person to the processing of their health data is discussed in Articles 6 and 9. Articles 21 and 22 address concerns about discrimination. To protect a person's privacy, Article 4 identifies which types of information should be kept private. I will show that none of these principles can protect a person from the harms that arise when they are the target of pandemic surveillance. These suggest that a specific regulatory framework be developed, focusing on safeguarding information attributed to a person because they are a member of a particular group.

The cluster-type (or statistical) groupings designed by big data analytics are sources of information for making policy decisions without focusing on individual identifiability. Regarding this, obligations or regulations developed to protect individuals from the misuse of their data are not helpful at the level of the

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

group, as groups created by algorithms or models expose those groups to potential harms without identifying individuals. Furthermore, current rules or regulations cannot protect groups against potential harms, partly because they focus on individual data protection concerns, and partly because many of the uses of big data that involve algorithmic groupings are so beneficial in furthering scientific research and improving public health. These suggest that while there are rules and obligations at the level of the individual, we must reach a stage in the development of data analytics where groups are protected against discrimination. The group should be granted privacy rights in order to limit the potential harms that can result from invasive and discriminatory data processing (Mantelero 2016). I will here investigate the feasibility of assigning group rights to the group clustered by big data analytics to mitigate harm to that group.

In the first part, I will look at the ethical issues raised by aggregate-level conclusions generated from big data that target people as members of groups, and groups *qua* groups. The second part will offer recommendations for how to improve current safeguards for persons as members of specific groups and for groups as a whole.

Key Ethical Issues

In this section, I will first look at ethical issues raised by aggregate-level conclusions generated by or discovered from big data while targeting a person as a member of a group. Three of the ethical issues, consent, social justice and fairness, and privacy, will be discussed in this section. Second, I will look at ethical issues raised by the targeting of a group *qua* group. Group discrimination or stigmatization will be discussed in this section. I acknowledge that there are other ethical issues not listed here, and so this list is not intended to be exhaustive. However, it covers the major issues that arise in the literature.

Ethical Concerns Raised by the Targeting of a Person as a Member of a Group

This section will deal with ethical issues that arise due to a person being targeted as a member of a group. To approach this, I will first provide a brief overview of the various types of groups created by data technologies. The distinction between different groups will enable a clearer explanation of the ethical issues. Data technologies are used to discover new patterns and relations in data. Those patterns and relations may concern numerous entities leading to profiles being formed, which in this context would be profiles of people. A profile which is a property or collection of properties of a particular group of people is known as a

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

group profile. Group profiles are divided into two types concerning the distributivity of properties forming group profiles. First, if a property is valid for each individual member of a group, this is called distributivity or a distributive property. Second, when a property is valid for the group and for individuals as members of that group, though not for those individuals as such, this is called non-distributivity or a non-distributive property (Vedder 2000).

Distributive generalizations and profiles attribute properties to a person, or a group of people, in such a way that these properties are actually and unconditionally manifested by all members of that group. For example, having a bad health condition may be distributed among all members of a group (those who have that condition). Non-distributive generalizations and profiles, on the other hand, are framed in terms of probabilities, averages, and medians, or significant deviations from other groups. They are based on comparisons of group members with one another and/or comparisons of one group with other groups. As a result, non-distributive generalizations and profiles differ significantly from distributive generalizations and profiles. Non-distributive generalizations and profiles apply to people as members of the reference group, but these individuals do not have to display these properties in reality (Vedder 1999). For example, in epidemic research, a property may be assigned to a patient because the person belongs to a reference group, such as having a specific disease, which is non-distributive profile information, even when the patient does not get sick from the disease. In such a circumstance, the person being judged and treated is being judged and treated on the basis of belonging to the 'wrong' category of persons.

In a distributive group profile, each individual member of the group is examined, the property discovered is assigned to each member, and the group inherits the property. For example, each patient in a group is diagnosed with a certain disease based on the presence of a certain symptom, and the property is then assigned to the entire group. We can conclude that the group inherits a distributive profile shared between all members of the group. However, in a non-distributive profile, the pattern or property discovered in a group is only distributed among parts of the group. In such cases, though, the property is ascribed to each member of a group because they are the members of the group (Vedder 2000), and not because they necessarily have that property. As a result, while the probabilistic property is ascribed to the group, attribution of that property to each and every individual member is invalid because that property may or may not ascribe to a particular person in the group. For example, when a group profile states that 90 per cent of the patients in the group have a particular symptom, no one can tell on those data alone, which patients actually do have the symptom. The link connecting the non-distributive profile to the individual to whom the group profiles may apply is opaque. Hence, this type of group profiling represents a group and reveals attributes that may (or may not) be applicable to the individuals in the group, and is only applicable to the group as such (de

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

Andrade 2011). Thus, assigning a non-distributive group profile to a group does not imply assigning that property to each of the group's members, implying that a group and its members do not share the same property.

I can now turn to the ethical issues that arise when a person is targeted using information derived from big data analytics.

Consent

Consent has been a point of debate and concern since its position of dominance in the post–Second World War Nuremberg Code, a set of ethical principles for human experimentation to ensure that harms to humanity like those in Nazi 'medical' experiments would never occur again (Annas and Grodin 1995; Macnish 2019). The purpose and justification of consent provisions are to provide reasonable assurance that a patient or research subject has not been deceived or coerced (O'Neill 2003). Hence, when research is aimed at impacting the conditions of its subjects, it is necessary to pay attention to research subjects' consent and awareness.

The function of consent in the big data era should be to help reduce harms associated with targeting members of a specific group. An example of potential harm perceived on group membership and not on individuals is tracking migrants fleeing a capital city in order to target cholera prevention measures (Bengtsson et al. 2011) through restriction of their travel. In this case, the question arises of how to manage big data sources in terms of consent and awareness among research subjects—as members of a specific group. To gain a better understanding of the issue, consider how group profiles are designed once more. Big data analytics are used to design group profiles to help control disease outbreaks, which are often based on fluid and contingent factors such as postal code, health status, and being in a public place at a specific time. In such cases, groups are not stable but fluid, and they are not unique or sparse but rather omnipresent and widespread. Group profiles can be designed in a fraction of a second and changed by changing the purpose and needs of grouping individuals in a specific way, so who is in and who is out of a group profile can change frequently (Floridi 2017; Sloot 2017). Thus, the issue is how to seek and obtain consent when members of a group may be unaware that they are part of a group and are included in a group because they share characteristics such as being in the same place at the same time.

In the context of big data analytics, there are two main limits to obtaining consent from those who are surveilled and grouped in a specific way. First, due to the unforeseen inferences drawn from data analytics, the possible risks and benefits might not be anticipated or anticipable at the time of initial data collection. Second, the problem stems from an inability to provide individuals with the option to choose which types of groups they want to be a part of and then make group decisions based on that. While novel approaches to consent are being

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

developed (e.g., dynamic consent, open consent, e-consent (Budin-Ljøsne et al. 2017; Kaye et al. 2015)), there is still a lack of giving individuals the choice to decide whether to be a member of a specific group simply because they share characteristics with other members of an algorithm-designed group.

Social Justice and Fairness

Group profiles, in this context, are designed and used only for pandemic research purposes, with guarantees that access to them is restricted to some researchers who do not share the information with others. However, things change when these guarantees are not present. The information in the profiles may then be made available to others, becoming part of the body of public knowledge in society, or the information may be used for entirely different purposes. For example, the information generated from people's health data could be used for other purposes and by third parties: for job selection procedures, insurance, loans, determining who can and cannot get back to work, or determining who can and cannot access public spaces like subways, malls, and markets (Morley et al. 2020; Sharon 2020; Vedder 2000). If this type of mission creep (Mariner 2007) occurs, then values of social justice and fairness are at stake.

Firstly, when the allocation of goods and amenities in society is based on health criteria, social justice is at issue. Generalizations and profiles can be used to help public and private entities formulate policies, or they can be absorbed into public knowledge. When the information contained in the generalizations or profiles is sensitive in nature, the situation becomes more complex because it might render members of the group vulnerable to prejudice or it may be used to make decisions regarding the allocation of scarce welfare resources. Information about people who have a high risk of developing certain diseases, especially those which may indicate a likely lifestyle, for example, can lead to stigmatization and prejudice. This information might be used to provide or restrict access to services such as insurance, loans, or jobs for members of a specific group. As a result, social justice challenges arise from some of the policy reactions to the information discovered from group profiles (Vedder 2000).

Secondly, fairness is at stake because an individual may be judged or treated based on merits or characteristics that he or she did not acquire voluntarily, such as a poor health condition. However, because the feature is one of the group and not necessarily of the individual, a person as such may not exhibit or even experience those characteristics at all. This occurs when non-distributive generalizations and profiles are used instead of distributive generalizations and profiles.

Privacy

Data technologies are used to find patterns or relationships in a dataset through maximizing dissimilarities between groups and optimizing similarity within a group (Aouad, Le-Khac, and Kechadi 2007). As mentioned above, the patterns

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

or relationships uncovered could apply to various entities, resulting in the formation of individual or group profiles. Group profiles may be used to infer characteristics to individuals (Henschke 2017). For example, the aggregation of data may result in the knowledge that those with low oxygen saturation may be more likely to be infected with Covid-19. Thus, algorithms design a group with low oxygen saturation, which is labelled as having a high risk of infection. Consider the case where a person's data were collected, stored, and processed, and the information 'low oxygen saturation' is attributed to him or her. This information might help clinicians make early decisions regarding the arrangement and organization of medical resources and early interventions to improve the health outcomes of this patient (Benito-León et al. 2021).

However, inferring group characteristics to individuals threatens the privacy of the individual as a member of a group. Inferred information tells us something about individual members of those groups in a very qualified way (Vedder 2000), assuming that the information is produced in a sound and reliable way. When an individual member intends to keep that information private, or when the information inferred is contrary to an individual member's preference, the privacy of members, rather than individual privacy, is threatened. The reason for this is that issues of individual privacy arise when the information generated is uniquely about a specific individual, meaning that the link between that individual and the information generated is strong. However, there are privacy issues when the link between the information generated and that individual is weak, especially in a non-distributive group profile, meaning that the information produced could have been formed from another source. In such cases, privacy claims are derived from group claims following the aggregation of the data (Henschke 2017). As a result, given the lack of direct connection to the individual source, inferring group characteristics to individuals in situations where a person is a data source threatens the privacy of members, implying that a more in-depth examination of how the privacy of groups' members is considered in the context of data protection is required.

Ethical Concerns Raised by Targeting a Group *Qua* Group

In this section, I will look at group discrimination or stigmatization when a group is targeted. Consider an epidemic that appears to target certain minorities disproportionately, resulting in additional restrictions being imposed on those minority groups, regardless of whether members of the group have the disease. In what follows, group discrimination or stigmatization will be discussed.

Contact tracing apps, GPS ankle monitors and other wearables, cell phone location data collection, genomic testing, and targeted quarantines, among other bio-surveillance technologies being used to respond to the Covid-19 pandemic, have the potential to exacerbate discrimination against racial minorities and

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

immigrants. As a result of the Covid-19 pandemic, racial disparities in health outcomes have increased, while communities of colour, immigrants, and other marginalized groups have been blamed for spreading the disease. Disturbing disparities in Covid-19 surveillance of racial minorities have emerged, for example, in the United States. In New York City, Black or Latinx people made up 92 per cent of those arrested for violating Covid-19 protocols, such as social-distancing requirements. Black people were targeted by government authorities at four and a half times the rate of White people for such violations (Sundquist 2021). As a result of 'inappropriate surveillance' (Macnish 2012), certain population groups, namely immigrants and certain non-White racial groups, are discriminated against and blamed for disease outbreaks, which may represent a biased evaluation and become a source of social discrimination.

Making inferences and drawing conclusions about groups based on an extensive collection of information threatens the group's privacy because revealing this information increases the risk of potential harm to the group itself. Hence, the surveillance technologies used in the fight against Covid-19 have an impact on the privacy of some groups, such as marginalized communities. That is, even if all members of a marginalized group are individually protected from unwanted intrusion and targeting, the group as a whole is not protected against disproportionate surveillance, implying that individual privacy can be effectively protected while the group as a whole is not adequately protected.

Consider a situation in which each individual knowingly shared his or her data and agreed to the type of processing to be performed at the time. Assume that the lawfully obtained and lawfully processed set of personal data enabled an analyst to draw sophisticated inferences—say, on the likelihood of disease outbreaks among populations—predicting the behaviour of a group of individual data subjects as a group. Such inferences would be based not on analysing past individual behaviour to predict future individual behaviour, but rather on comparing and contrasting the behaviours of all members of a group defined by one or more shared characteristic (Kammourieh et al. 2017). Disclosing information discovered about a group therefore increases the risk of harm to that group's privacy because it increases the risk of discrimination against the group.

Current Measures to Address the Identified Issues

In this section, I will look at the current guiding regulation regarding data protection, the General Data Protection Regulations (GDPR), to explore the suitability of existing legal frameworks to address and mitigate the identified issues. I demonstrate that further work is required to address the identified issues and that specific rules or regulations need to be developed that differ from those already existing regulations in the field of data protection.

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

Protecting Persons against Harms

Article 9 provides the legal ground for special categories of personal data in the context of epidemics. Processing of special categories of personal data, such as health data 'for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health' is allowed. These special categories of personal data are processed for reasons of public interest without the consent of the data subject. This Article is unable to address the identified consent issues and instead introduces a new issue in the form of the privacy–health trade-off.

Profiling and discrimination concerns are reflected in GDPR, especially in Article 21. This Article introduces the right of data subjects to object to personal data processing, including profiling, at any time. If the purpose of data processing is direct marketing, the data subject will have an absolute right to object (Wachter 2018). However, the scope of the Article is limited to individual profiles that analyse or predict specific aspects of natural persons without taking into account harms that arise when a person is considered as a part of a whole group, particularly non-distributive group profiles in which the analysis or prediction is performed by comparing and contrasting the behaviour of all members of a group, rather than predicting behaviour of a specific person based on his or her available data.

Article 4(1) determines which types of information are protected by GDPR. Personal data allowing for identification of a natural person, including online identifiers or factors specific to the physical, physiological, genetic, mental, economic, cultural, or social identity of that natural person, are protected. Nevertheless, as soon as the data have ceased to be personal data in the strict sense, it is not clear how the principle should be applied. For example, the right of controlling data does not apply to information derived from personal data (Vedder 2000). As a result, in the age of big data and information inferred, the interest in informational privacy no longer provides sufficient protection to the *individual members of a group*; it focuses solely on information collection rather than analysis of *aggregation data* (Kammourieh et al. 2017).

In order to address the issues, we need to rethink and expand our current moral vocabulary and legal frameworks for dealing with information technology. Broadening the scope of information protected by the right to privacy and data protection to include information primarily attributed to a person because of their membership in a specific group is one way to address the shortcomings of current privacy conceptions in relation to big data analytics (for more information, see Vedder's definition of categorical privacy (1999)). Furthermore, Henschke (2017) and Kammourieh et al. (2017) propose the protection of metadata, the valuable information that can be inferred from datasets, rather than raw data, as a way to address privacy issues.

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

Protecting Groups against Harms

According to Mantelero (2016), group privacy is the right to limit the potential harms to the group itself that can derive from invasive and discriminatory data processing.[1] At the group level, the right to privacy can be perceived as a duty of the state not to use its powers arbitrarily. A group right to privacy prevents the arbitrary use of power, such as discriminating illegitimately between different groups in society or exercising power for no reason at all (Sloot 2017). Understanding group privacy in terms of protecting groups against the possible negative consequences of generalizations and profiles cannot be reduced to individual privacy, meaning that the protection of group members cannot protect the group itself.

It could be asserted that, in some cases, the protection of individuals can protect specific groups. GDPR, for example, has the potential to provide safeguards against groups. GDPR provides enhanced protection for certain types of highly sensitive data, including 'revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, and [. . .] the processing of data concerning health or sex life' (Art. 9, EU Parliament 2016). While this is a protection granted to the individual, its effect is also to protect specific groups that are more vulnerable to targeting (Kammourieh et al. 2017). As a result, GDPR has the potential to limit the discovery of information about existing groups, such as racial groups.

However, GDPR is mainly focused on protecting individual identity and on safeguarding personal information. In an era of big data, where information about groups is extracted from data, or where more information is discovered about existing groups, the individual is often incidental to the analysis. Thus, the problem is not that this or that specific person has been affected, but that groups have been harmed. Since the group is exposed to the risks derived from the creation and use of inferred data, the infringement takes place at the group level while the rights and remedies are granted at the individual level (Kammourieh et al. 2017; Taylor, Floridi, and Sloot 2017; Sloot 2017). Regarding this, it is important to assign rights to a group to protect that group against discriminative harms. Granting this right to groups is different from the existing right in the field of privacy and data protection, in that this right to privacy is not reducible to the privacy of the individuals forming such groups.

There are at least two reasons why group rights to privacy cannot be reduced to individual rights to privacy. First, the training set used to develop a model and then generate an inference may not include all members of the group (e.g., patients with a specific disease), implying that, while there is no violation of the individual right (because members of the training set provided full informed

[1] It is assumed that the group right to privacy is limited to the right against discrimination.

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

consent to the collection, analysis, and inference of the data), there is a violation of the right to the broader group, i.e. (the set of all patients with the same disease) minus (the set of people of the same disease in the training set). Second, the information discovered by big data analytics may be targeted at specific groups comprising different individuals with diverse interests. Due to this, the group's interests[2] should fulfil the individual interests of diverse members at the same time: those who have no interest in limiting their information usage and those who do. Consequently, the group's interests may not correspond to the interests of each individual member, meaning that the group's interests are not the result of the aggregation of its members' individual interests. Irreducibility of the group's interests to a simple aggregation of individual interests implies that the group's right to privacy must be invoked to protect non-aggregative group interests.

The preceding discussions highlight the importance of developing group rights to privacy to address issues that revolve around the risks of discrimination and the adverse outcomes of big data analysis. In what follows, I will discuss the feasibility and problems of ascribing group rights to privacy to a clustered group.

Group Rights to Privacy

So far, I have discussed that, in contrast to how privacy has traditionally been conceived on an individual level, the era of big data raises new questions about privacy on a group level. In such cases, access to personally identifiable information of individuals is less likely to cause harm. Harm is more likely to occur when authorities or corporations draw inferences about people on a group level. As a result, the concept of privacy must be stretched and reshaped in order to help us think about groups. Floridi (2014) was the first to bring up the concept of group privacy in relation to big data analytics insights. He argued that it is crucial to investigate whether groups have privacy rights that are not reducible to the privacy of the individuals who make up those groups.

According to Floridi's argument, a group's right to privacy is a right held by a group *qua* group rather than its member severally; it is referred to as a group right in the *strong* sense, the corporate approach to group rights. Right-holding groups are conceived as moral entities in their own right, with a being and status similar to that of an individual person. This viewpoint holds that a group has an identity and existence distinct from its members. Accordingly, *unity* and *identity* are necessary for a group to be the type of group that can bear rights (French 1984; Newman 2011; Taylor et al. 2017; Taylor & Floridi 2017). For example, French

[2] According to the choice (will) theory of rights, to have a right is to have a choice, so it makes sense to ascribe rights only to beings who are capable of choice (Hart 1982). Since clustered groups clearly lack the capability of choice, I limit myself to the interest (benefit) theory of rights.

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

(1984) contends that the Gulf Oil Corporation's rights and responsibilities in purchasing or selling property, or in being responsible for environmental pollution and cleaning it up, are not reducible to the individuals currently associated with it. Organizations of this type have identities that are not exhausted by the identities of the people who work in them; one person leaving and another joining does not form a new organization. As a result, a group's unity or identity distinguishes it as the type of group that might have rights.

However, proponents of group rights in the *moderate* sense, the collective approach to group rights, such as Raz (1988), argue that groups are not conceived as having independent standing, but rather as having rights shared in and held jointly by the group's members, rather than being a mere aggregation of rights held by the group's members individually. The individuals who comprise the group have a right that none of them have as independent individuals. In this view, collective rights are ascribed to a specific collection of individuals because there are some sorts of public goods that can only be held by the collective. In respect of the public production of such goods, participants in a participatory activity possess collective rights. For example, the provision of a cultured society requires participation amongst members of the group; each individual needs others in order to produce the desired society. Accordingly, there is no individual right to a cultured society, but rather participants in a joint action possess collective rights (Miller 2001; Raz 1988).

I argue that, in the case of a group designed by algorithms or data technologies, it is implausible to regard a group's right to privacy as a group right in either the strong or moderate sense. The reason for this is that in this kind of group, the essential criteria for both strong and moderate approaches on group rights to privacy are not met. On the one hand, because of the lack of integrity or unity needed to hold a right according to strong approaches, a group's right to privacy cannot be described based on these approaches. A group's right to privacy, on the other hand, cannot be conceived based on moderate approaches because members of the relevant group cannot perform a joint action to produce any good simply because they cannot realize the condition required to constitute a joint action, which is 'believing that their action is dependent on the action of other members' (Miller 2001: 57).

Although a group's right to privacy (for a group designed by algorithms) cannot be explained theoretically using either strong or moderate approaches to group rights, we can take methodological approaches to justify why a cluster-designed group requires such a right. For this, I propose employing constructivist theories implying that the need for a moral group's right to privacy is practical. According to such theories, any reason that justifies a right as a moral right must be morally neutral (Copp 1995). Thus, the justification of a moral right must be explained by invoking non-moral values. From this perspective, I claim that a group designed by algorithms would be rationally required to have a group's moral right to privacy to meet its non-moral values, if any, associated with such a right. For example, we need

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

to grant a right to privacy to a clustered group to protect health (or sensitive) information about that group in cases where revealing or releasing that information about the group affects the members' relationships with others in society.

Conclusion and Recommendations to Improve Current Measures

Big data analytics have the capacity to uncover new information, find patterns, and predict behaviour, allowing for the algorithmic creation of totally new groups. In this regard, it is necessary to reconceptualize the risk of data harm to include the problem of the group and its members. For researchers, it is difficult to manage the source of big data regarding consent and awareness on the part of research subjects. A further problem is that the application of data technologies undermines the values of social justice and fairness, since an individual may be judged or treated on characteristics they did not acquire voluntarily (or at all). Finally, because data technologies are used to target people as members of specific groups rather than individuals, they are increasingly threatening group members' privacy rather than individual privacy. In addition to the issues that arise when a person as a member of a specific group is targeted by the information derived from a group profile, there are also risks to the privacy of a group *qua* group because revealing the information about a group increases the risk of discrimination against that group.

In order to protect groups as such, I agree with Floridi (2014) that clustered groups must have rights to privacy which do not reduce to the privacy of individuals forming such groups. I also agree with Mantelero (2016) that a group right to privacy is required to limit the potential harms that can result from invasive and discriminatory data processing. However, group rights to privacy cannot be theoretically ascribed to a clustered group using traditional approaches. In terms of the significance of granting such rights, I recommend taking a methodological rather than predominant standard approaches to interpret moral group rights to privacy.[3]

References

Andrade, Norberto Nuno Gomes de. 2011. 'Data Protection, Privacy and Identity: Distinguishing Concepts and Articulating Rights'. In *Privacy and Identity Management for Life*, edited by Simone Fischer-Hübner, Penny Duquenoy, Marit

[3] Haleh Asgarinia is supported by the PhD fellowship from the 'PROTECT—Protecting Personal Data Amidst Big Data Innovation' project, funded by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 813497.

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

Hansen, Ronald Leenes, and Ge Zhang, 352:90–107. IFIP Advances in Information and Communication Technology. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-20769-3_8.

Annas, George J., and Michael A. Grodin, eds. 1995. *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*. New York: Oxford University Press.

Aouad, Lamine M., N. Le-Khac, and M. Kechadi. 2007. 'Lightweight Clustering Technique for Distributed Data Mining Applications'. In *Industrial Conference on Data Mining*. https://doi.org/10.1007/978-3-540-73435-2_10.

Bengtsson, Linus, Xin Lu, Anna Thorson, Richard Garfieldarfier, and Johan Schreeb. 2011. 'Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti'. *PLoS Medicine* 8 (August): e1001083. https://doi.org/10.1371/journal.pmed.1001083.

Benito-León, Julián, Ma Dolores del Castillo, Alberto Estirado, Ritwik Ghosh, Souvik Dubey, and J. Ignacio Serrano. 2021. 'Using Unsupervised Machine Learning to Identify Age- and Sex-Independent Severity Subgroups Among Patients with Covid-19: Observational Longitudinal Study'. *Journal of Medical Internet Research* 23 (5): e25988. https://doi.org/10.2196/25988.

Budin-Ljøsne, Isabelle, Harriet J. A. Teare, Jane Kaye, Stephan Beck, Heidi Beate Bentzen, Luciana Caenazzo, Clive Collett, et al. 2017. 'Dynamic Consent: A Potential Solution to Some of the Challenges of Modern Biomedical Research'. *BMC Medical Ethics* 18 (1): 4. https://doi.org/10.1186/s12910-016-0162-9.

Chang, Serina, Emma Pierson, Pang Wei Koh, Jaline Gerardin, Beth Redbird, David Grusky, and Jure Leskovec. 2021. 'Mobility Network Models of Covid-19 Explain Inequities and Inform Reopening'. *Nature* 589 (7840): 82–7. https://doi.org/10.1038/s41586-020-2923-3.

Copp, David. 1995. *Morality, Normativity, and Society*. Oxford: Oxford University Press.

EU Parliament. 2016. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA Relevance). OJ L.* Vol. 119. http://data.europa.eu/eli/reg/2016/679/oj/eng.

Floridi, Luciano. 2014. 'Open Data, Data Protection, and Group Privacy'. *Philosophy & Technology* 27 (1): 1–3. https://doi.org/10.1007/s13347-014-0157-8.

Floridi, Luciano. 2017. 'Group Privacy: A Defence and an Interpretation'. In, 83–100. https://doi.org/10.1007/978-3-319-46608-8_5.

French, Peter A. 1984. *Collective and Corporate Responsibility*. New York: Columbia University Press.

Hart, H. L. A. 1982. 'Legal Rights'. In *Essays on Bentham*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198254683.003.0008.

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

Henschke, Adam. 2017. *Ethics in an Age of Surveillance: Personal Information and Virtual Identities*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316417249.

Kammourieh, Lanah, T. Baar, J. Berens, E. Letouzé, Julia Manske, J. Palmer, David Sangokoya, and P. Vinck. 2017. 'Group Privacy in the Age of Big Data'. In. https://doi.org/10.1007/978-3-319-46608-8_3.

Kaye, Jane, Edgar A. Whitley, David Lund, Michael Morrison, Harriet Teare, and Karen Melham. 2015. 'Dynamic Consent: A Patient Interface for Twenty-First Century Research Networks'. *European Journal of Human Genetics* 23 (2): 141–6. https://doi.org/10.1038/ejhg.2014.71.

Liu, Jiming, Bo Yang, William K Cheung, and Guojing Yang. 2012. 'Malaria Transmission Modelling: A Network Perspective'. *Infectious Diseases of Poverty* 1 (November): 11. https://doi.org/10.1186/2049-9957-1-11.

Macnish, Kevin. 2012. 'Unblinking Eyes: The Ethics of Automating Surveillance'. *Ethics and Information Technology* 14 (2): 151–67. https://doi.org/10.1007/s10676-012-9291-0.

Macnish, Kevin. 2019. 'Informed Consent'. In *Data, Privacy and the Individual*, edited by Carissa Veliz. Madrid: IE University Press.

Mantelero, Alessandro. 2016. 'From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era'. In, 139–58. https://doi.org/10.1007/978-3-319-46608-8_8.

Mariner, Wendy K. 2007. 'Mission Creep: Public Health Surveillance and Medical Privacy'. SSRN Scholarly Paper ID 1033528. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=1033528.

Mikkelsen, Daniel, Henning Soller, and Malin Strandell-Jansson. 2020. 'Data Privacy in the Pandemic | McKinsey'. 2020. https://www.mckinsey.com/business-functions/risk-and-resilience/our-insights/privacy-security-and-public-health-in-a-pandemic-year. *Account*. Cambridge: Cambridge

Miller, Seumas. 2001. *Social Action: A Teleological Account*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511612954.

Morley, Jessica, Josh Cowls, Mariarosaria Taddeo, and Luciano Floridi. 2020. 'Ethical Guidelines for Covid-19 Tracing Apps'. *Nature* 582 (7810): 29–31. https://doi.org/10.1038/d41586-020-01578-0.

Newlands, Gemma, Christoph Lutz, Aurelia Tamò-Larrieux, Eduard Fosch Villaronga, Rehana Harasgama, and Gil Scheitlin. 2020. 'Innovation under Pressure: Implications for Data Privacy during the Covid-19 Pandemic'. *Big Data & Society* 7 (2): 2053951720976680. https://doi.org/10.1177/2053951720976680.

Newman, Dwight. 2011. *Community and Collective Rights: A Theoretical Framework for Rights Held by Groups*. 1st edition. Hart Publishing.

O'Neill, O. 2003. 'Some Limits of Informed Consent'. *Journal of Medical Ethics* 29 (1): 4–7. https://doi.org/10.1136/jme.29.1.4.

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024

Raz, Joseph. 1988. *The Morality of Freedom*. Clarendon Paperbacks. Oxford: Oxford University Press. https://doi.org/10.1093/0198248075.001.0001.

Sharon, Tamar. 2020. 'Blind-Sided by Privacy? Digital Contact Tracing, the Apple/ Google API and Big Tech's Newfound Role as Global Health Policy Makers'. *Ethics and Information Technology*, July. https://doi.org/10.1007/s10676-020-09547-x.

Sloot, Bart van der. 2017. 'Do Groups Have a Right to Protect Their Group Interest in Privacy and Should They? Peeling the Onion of Rights and Interests Protected Under Article 8 ECHR'. In, 197–224. https://doi.org/10.1007/978-3-319-46608-8_11.

Sundquist, Christian Powell. 2021. 'Pandemic Surveillance Discrimination'. *SETON HALL LAW REVIEW* 51: 13.

Tatem, A., Y. Qiu, David L. Smith, O. Sabot, Abdullah S. Ali, and Bruno Moonen. 2009. 'The Use of Mobile Phone Data for the Estimation of the Travel Patterns and Imported Plasmodium Falciparum Rates among Zanzibar Residents'. *Malaria Journal*. https://doi.org/10.1186/1475-2875-8-287.

Taylor, Linnet. 2016. 'Safety in Numbers? Group Privacy and Big Data Analytics in the Developing World'. SSRN Scholarly Paper ID 2848825. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=2848825.

Taylor, Linnet, and Luciano Floridi. 2017. 'Group Privacy: New Challenges of Data Technologies'. *Group Privacy*, 293.

Taylor, Linnet, Luciano Floridi, and Bart van der Sloot. 2017. 'Introduction: A New Perspective on Privacy'. In *Group Privacy: New Challenges of Data Technologies*, edited by Linnet Taylor, Luciano Floridi, and Bart van der Sloot, 10–22. Philosophical Studies Series. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-46608-8_1.

Vedder, Anton. 2000. *Law and MedicineCurrent Legal Issues Volume 3*. Edited by Michael Freeman and Andrew Lewis. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198299189.001.0001.

Vedder, Anton. 1999. 'KDD: The Challenge to Individualism'. *Ethics and Information Technology* 1 (4): 275–81. https://doi.org/10.1023/A:1010016102284.

Wachter, Sandra. 2018. "Normative Challenges of Identification in the Internet of Things: Privacy, Profiling, Discrimination, and the GDPR." *Computer Law & Security Review* 34 (3): 436–49. https://doi.org/10.1016/j.clsr.2018.02.002.

Downloaded from https://academic.oup.com/book/55101/chapter/423908063 by guest on 25 January 2024