Philos. Technol. (2014) 27:1–3
DOI 10.1007/s13347-014-0157-8

EDITOR LETTER

Open Data, Data Protection, and Group Privacy

Luciano Floridi

Published online: 18 February 2014

The debate on Open Data and Data Protection focuses on individual privacy. How can the latter be protected while taking advantage of the enormous potentialities offered by ever-bigger Open Data and ever-smarter algorithms and applications? The tension is sometimes presented as being asymmetric: between the *ethics* of privacy and the *politics* of security. In fact, it is ultimately ethical. Two moral duties need to be reconciled: fostering human rights and improving human welfare. The tension is obvious if one considers medical contexts and biomedical Big Data, for example, where protection of patients' records and cure or prevention of diseases need to go hand in hand (Howe et al. 2008; Groves et al. 2013).

Currently, the balance between these two moral duties is implicitly understood within a classic ontological framework. The beneficiaries of the exercise of the two moral duties are the individual vs. the society to which the individual belongs. At first sight, this may seem unproblematic. We work on the assumption that these are the only two "weights" on the two sides of the scale. Such a framework is not mistaken, but it is dangerously reductive, and it should be expanded urgently. For there is a third "weight" that must be taken into account by Data Protection, that of groups and their privacy. Privacy as a group right is a right held by a group as a group rather than by its members severally. It is the group, not its members, that is correctly identified as the right-holder. A typical example is the right of self-determination, which is held by a nation as a whole.

The idea that groups may have a right to privacy is not new (see for example Bloustein (1978, 2003)) and it is open to debate (Bisaz 2012). But it has not received the attention it deserves, although it is becoming increasingly important. This because, by far, most people are not targeted by ICTs as individuals but as members of specific groups, where the groups are the really interesting focus, as carriers of rights, values, and potential risks. Think of owners of such and such kind of car, shoppers of such and such kind of food, people who like this kind of music, or people who go to that kind of restaurant, cats owners, dogs owners, people who live at that kind of postal code address, carriers of a specific gene, people affected by a certain disease, … Open Data is more likely to treat types (of customers, users, citizens, demographic population, etc.)

L. Floridi (✉)
Oxford Internet Institute, University of Oxford, 1 St Giles, Oxford OX1 3JS, UK
e-mail: luciano.floridi@oii.ox.ac.uk

rather than tokens (you, Alice, me…), and hence groups rather than individuals. But re-identifiable groups are *ipso facto* targetable groups. It is therefore a very dangerous fallacy to think that if we protect personal data that identify individuals, the protection of the groups will take care of itself.

Such an "atomistic" ontology—take care of each member separately and the group will automatically be fine too—is at the roots of current European legislation. This defines a "Data Subject" as:

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

As a consequence, both the 1995 Directive and the new Directive under discussion focus on individual persons. Here are the two texts for comparison (emphases added):

> Whereas the principles of protection must apply to any information concerning an identified or identifiable *person*; whereas, to determine whether a *person* is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said *person*; whereas the principles of protection shall not apply to data rendered anonymous in such a way that the *data subject* is no longer identifiable […].(Art. 26, Directive 95/46/EC)

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

> The principles of protection should apply to any information concerning an identified or identifiable *natural* person. To determine whether a *natural* person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the individual. The principles of data protection should not apply to data rendered anonymous in such a way that the data subject is no longer identifiable. (Art. 16, COM(2012) 10 Final 2012/0010 (COD)).

What we should acknowledge is the fact that both friendly and hostile users of Open Data may not care about Alice at all, but only about the fact whether some people, whoever they are, belong to the group that regularly goes to the local church, or mosque, or synagogue, uses Grindr, or has gone to a hospital licensed to carry out abortions, or indeed shares a feature of your choice. In military terminology, Alice is hardly ever a High Value Target, like a special and unique building. She is usually part of a High Pay-off Target, like a tank in a column of tanks.

As I have argued elsewhere (Floridi 2013) our current ethical approach is too anthropocentric (only natural persons count) and atomistic (only the single individual count). We need to be more inclusive because we are underestimating the risks involved in opening anonymised personal data to public use, in cases in which *groups* of people may still be easily identified and targeted. Such inclusiveness should not be

too hard to achieve. After all, we already accept as ordinary the fact that groups as agents may infringe on someone's privacy. In the USA, we are used to consider as normal collective lawsuits (class actions) in which a group may sue a person or another group. And in Europe, consumer organisations regularly bring claims on behalf of the groups they represent. Clearly, there are cases in which the protection of a right requires a balance between the agents issuing the action and the patients receiving the action.

There are very few Moby-Dicks. Most of us are sardines. The individual sardine may believe that the encircling net is trying to catch it. It is not. It is trying to catch the whole shoal. It is therefore the shoal that needs to be protected, if the sardine is to be saved. An ethics addressing each of us as if we were all special Moby-Dicks may be flattering and it is not mistaken, but needs to be upgraded urgently. Sometimes the only way to protect the individual is to protect the group to which the individual belongs. Preferably before any disaster happens.

Acknowledgements I am very grateful to Massimo Durante for his most valuable comments on a previous draft of this article; to Ugo Pagallo and the participants in the panel "Open Data and Data Protection: Problems and Perspectives"—organised at the conference *Computers, Privacy and Data Protection 2014 (CPDP 2014) Reforming Data Protection: The Global Perspective*—for their feedback; and to Linnet Taylor for some enlightening conversations on the topic of group privacy. Her article (Taylor 2014) makes a strong and convincing case for the protection of group privacy in contexts of geolocated data, especially in Africa.

References

Bisaz, C. (2012). *The concept of group rights in international law : groups as contested right-holders, subjects and legal persons*. Leiden: Brill, Nijhoff.

Bloustein, E. J. (1978). *Individual and group privacy*. New Brunswick, N.J.: Transaction Publishers.

Bloustein, E. J. (2003). *Individual & group privacy* (2nd ed.). New Brunswick, N.J.: Transaction Publishers.

COM(2012) 10 Final 2012/0010 (COD). Proposal for a Directive of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data by competent authorities for the purposes of prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and the free movement of such data /* COM/2012/010 final - 2012/0010 (COD) */

Directive 95/46/EC. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

European Commission. Justice, Data Protection, Glossary, Data Subject: http://ec.europa.eu/justice/data-protection/glossary/index_en.htm.

Floridi, L. (2013). *The ethics of information*. Oxford: Oxford University Press.

Groves, P., Kayyali, B., Knott, D., & Van Kuiken, S. (2013). "The 'big data' revolution in healthcare." *McKinsey Quarterly*.

Howe, D., Costanzo, M., Fey, P., Gojobori, T., Hannick, L., Hide, W., et al. (2008). Big data: the future of biocuration. *Nature, 455*(7209), 47–50.

Taylor, L. (2014). "No place to hide? The ethics and analytics of tracking mobility using African mobile phone data." online version available at http://www.academia.edu/4785050/No_place_to_hide_The_ethics_and_analytics_of_tracking_mobility_using_African_mobile_phone_data. (in press)