Theory on Demand #29

Good Data

Editors: Angela Daly, S. Kate Devitt and Monique Mann

Copy editor: Harley Williams
Editorial assistant: Kayleigh Murphy

Funding: Queensland University of Technology Faculty of Law

Cover design: Katja van Stiphout
Design and EPUB development: Barbara Dubbeldam

Published by the Institute of Network Cultures, Amsterdam, 2019


Contact
Institute of Network Cultures
Phone: +3120 5951865
Email: info@networkcultures.org
Web: http://www.networkcultures.org

This publication is available through various print on demand services and freely downloadable from http://networkcultures.org/publications

This publication is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-SA 4.0).
17: ALGORITHMIC MAPMAKING IN 'SMART CITIES': DATA PROTECTION IMPACT ASSESSMENTS AS A MEANS OF PROTECTION FOR GROUPS

GERARD JAN RITSEMA VAN ECK

Introduction

Maps are powerful communication tools, and mapmaking used to be a privileged affair. In recent times, this has changed as ‘smart cities’ have been outfitted with video, audio, and other kinds of ‘Internet of Things’ sensing devices. The data-streams they generate can be combined with volunteered data to create a vast multitude of interactive maps on which individuals are constantly (re)grouped based on abnormality, deviation, and desirability. Many have argued that under these circumstances personal data protection rights should extend to groups.

However, group rights are an awkward fit for the current European Data Protection Framework, which is heavily focused on individuals. One possible opening for better protection is offered by Data Protection Impact Assessments (DPIAs), which are mandatory to carry out when the ‘systematic monitoring of a publicly accessible area on a large scale’ is necessary for mapmaking takes place. They form an opportunity to recognize the risks of e.g. discrimination at an early stage. Furthermore, by including representatives of local (disadvantaged) groups, the strong performative qualities of maps can offer occasions for groups of citizens in smart cities to proactively shape the environments in which they live.

There are serious limitations. Although DPIAs are mandatory, the inclusion of affected data subjects and their representatives is not. This undermines many of the possible advantages. Finally, the high costs associated with the process might mean many companies engage with it only superficially and temporarily. Establishing effective data protection for groups negatively impacted by mapmaking software through DPIAs thus seems nigh on impossible in lieu of substantial legislative change.

In late 2017, Strava revealed its ‘global heatmap’ which shows popular running and cycling routes around the world. Although first reported as a ‘striking visualization of over one billion activities’, the tone of the reporting quickly changed when it was discovered that secret (mili-

---

1 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 (GDPR)) (2016) OJ L119/1, point (c) of art 35(3).

tary) locations could also be located using the heatmap.\textsuperscript{3} The Stratumseind, a popular nightlife area in the Dutch city of Eindhoven received similarly negative press when it transformed into a ‘Living Lab’ where massive amounts of data about people’s activities will be used [...] to study which factors contribute to violence and discomfort.\textsuperscript{4} This data is overlaid on a map of the area, and the police can quickly intervene when and where suspicious patterns emerge. One reporter dubbed it ‘Having a beer with Big Brother’.\textsuperscript{5}

Clearly, the usage of large amounts of data to draw maps raises concerns. The practice, however, seems to be booming rather than decreasing. ‘Smart’ and dumb cities alike are increasingly being outfitted with video, audio, and various other ‘Internet of Things’ (IoT) sensing devices in order to gather ever more data,\textsuperscript{6} and an even more potent, and often cheaper, stream of data can be crowdsourced or scraped from (the smartphones of) city dwellers. Once generated, the data needs to go somewhere, and often they end up in maps.

This chapter will consider these developments by combining insights from critical geography and data protection law, particularly the General Data Protection Regulation (GDPR) that recently came into force in the European Union. Specifically, the chapter will begin by investigating the relationship between data and maps, and why they are such salient artifacts. In the next section, the issue of maps is problematized and the lack of legal protections for affected groups is discussed. Then, Data Protection Impact Assessments (DPIAs), a tool that has received renewed interest as a result of the introduction of the GDPR, are introduced as a possible solution in the absence of strong legislation. But are they really a panacea that can provide meaningful safeguards, or an expensive Band-Aid that will be avoided, ignored, and brushed aside?

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{heatmap.png}
\caption{Screenshot taken by the author on https://labs.strava.com/heatmaps on 2 February 2018.}
\end{figure}


Shown is the Stadspark in Groningen, the Netherlands. The clearly visible small semi-circle in the bottom is a 400-meter running track whereas the larger oval above it goes around a pond and measures almost exactly one kilometer; it therefore constitutes a popular running route.

How Maps Are Made

Drawing a map is a powerful act. Who draws a map, decides what is shown on it: physical features of the landscape, roads, settlements, or claims to certain territories by various (imagined) groups. Maps provide a representation of a certain geographical area that, even if not fully objective, at least is claimed to represent the truth. Map drawing has historically been a privileged affair that is only undertaken by those with the means and motivation to do so, including ‘scientists, planners, urbanists, technocratic subdividers and social engineers’. However, over the past two decades, the global north has seen a dizzying multiplication of geotagged databases visually overlaid on maps, each of which comprises a new map in itself. Such ‘Geographic Information Systems’, or simply GIS as they are referred to in the technocratic jargon, are often not carefully constructed by single authors but are based on aggregations of data from a multiplicity of sources, each of which might have been collected differently.

Outfitting an environment with enough IoT sensors to collect the critical mass of geotagged data needed to feed a GIS can be prohibitively expensive for all but the most well-funded smart cities. A cheap and therefore popular alternative, or addition, is crowdsourcing. Roughly, organizations that crowdsource the collection of georeferenced data in order to make maps have three options. One option is that data is collected purposefully by participants that have knowledge of (and perhaps an interest in) how it will be processed further. This is the approach taken by, for example, OpenStreetMap, which is ‘built by a community of map-makers that contribute and maintain data about roads, trails, cafés, railway stations, and much more’. Another option is to scrape together the various geotagged data trails that smartphone...
users leave behind, as is done by the Living Lab in Eindhoven and various others. The third option is to combine these two strategies. This model is successfully employed by the navigation app Waze, which takes into account the current traffic situation to calculate the fastest routes. To be able to do so, it scrapes data on the speed and location of its users and combines this with data on for instance road closures that users are encouraged to add using a gamified system of badges and achievements. Furthermore, in various cities such as Rio de Janeiro, New York City, and Ghent, Waze receives data on upcoming road closures from local governments through its Connected Citizens Programme - in exchange, the local governments gain access to information on accidents and other traffic disruptions. As Gekker and Hind point out, this makes it ‘not necessarily easy to make a clean split between those who “produce” the map, those who “edit” the map and those who “consume” the map’.

Figure 2. Screenshot taken by the author on https://www.waze.com/livemap on 5 February 2018. Shown is part of the center of London, England. The three data sources are clearly visible: the coloration of the streets is based on scraped movement data, the balloons are based on user submitted reports, and the map data comes from Google Maps.

Once collected, the manner in which crowdsourced georeferenced data is presented further blurs the lines between producers, editors, and consumers. Georeferenced datasets are

---

14 See e.g. Federico Botta, Helen Susannah Moat and Tobias Preis, ‘Quantifying Crowd Size with Mobile Phone and Twitter Data’, *Royal Society Open Science* 2.5 (May 2015), DOI: 10.1098/rsos.150162.
visualized as cartographic imagery in interactive environments overlaid on Google Maps or similar services like OpenStreetMaps or Bing Maps,\textsuperscript{17} engaging the user in acts of active sense making.\textsuperscript{18} This makes \textit{prima facie} sense, as the public is accustomed to engaging with large data sets visually.\textsuperscript{19}

Interestingly, something else is also gained when crowdsourced georeferenced data is offered as a map. This mode of presentation immediately embeds the data within what Leslie calls 'the "system of maps," the full panoply of mutually-reinforcing, mutually referential map images that subjects are exposed to [...], a system ultimately grounded in a generalized awareness of cartography’s scientism'.\textsuperscript{20} Thus, by becoming maps, crowdsourced georeferenced data sets also gain the performative power that comes with the claim to truth-correspondence of maps. Put differently: Because people tend to take maps at face value, any data overlaid is also taken at face value, whether it purports to show a crime hotspot or a traffic jam.

\textbf{Maps and (Missing) Group Rights}

The role of many modern algorithms is to render 'big data' actionable on the basis of hyper-complex probabilistics.\textsuperscript{21} They do this by categorizing people into groups: for example, those who should be shown a certain advertisement, those who should not be let into an airplane without an extra security check, or those who should receive special police attention because they know a victim of a recent shooting.\textsuperscript{22}

In public spaces, the push to make cities smart by outfitting them with a multitude of sensors also increases the influence of grouping algorithms.\textsuperscript{23} Take, for example, those who are unwittingly part of a suspicious pattern in the Stratumseind Living Lab in Eindhoven. Once identified, the location of the pattern is shown on a map, which, as outlined above, forms an effective communication method for the end-result of such algorithms: it embeds the abstract

\begin{enumerate}
\item Which itself also provides a seamless transition between multiple ‘views’ of the same area, such as ‘map,’ ‘terrain,’ ‘traffic,’ and ‘satellite.’
\item Leslie, ‘Territoriality, Map-Mindedness, and the Politics of Place’, 172.
\end{enumerate}
(and perhaps highly contestable)\textsuperscript{24} output of the algorithm in the ‘panoply of maps’.\textsuperscript{25} Private security personnel can then be deployed to re-establish order. The Strava example mentioned before shows that an environment does not need to be highly saturated with IoT sensors for such groupings to have an effect: only a few smartphones in the Syrian Desert were enough to give away the locations of various secret military bases.

The most problematic side product of map-making algorithms can be geographical biases. Although they do not directly affect individuals, they can deeply impact local communities\textsuperscript{26} through, for instance, increased police surveillance.\textsuperscript{27} This can, inter alia, be caused by a skew in the collection of crowdsourced geographic data, in which marginalized communities tend to be underrepresented,\textsuperscript{28} or conversely, their overrepresentation in historical data.\textsuperscript{29} After the collection phase, bias can also creep in during analysis. At this stage though, it may be harder to detect any biases\textsuperscript{30} because the inner workings of many algorithms can be difficult to understand for non-experts and, in the case of self-learning algorithms, for the developers of the algorithm itself.\textsuperscript{31} Such biases can easily turn into disparate impacts and discrimination when the output of an algorithm is taken at face value. This is especially salient for mapmaking algorithms because, as was discussed above, representing abstract outputs on maps increases their performativity.

The enigmatical nature of refined data processing means that even if strong anti-discrimination legislation exists, it might not provide much protection.\textsuperscript{32} This follows from the simple fact that,

\begin{itemize}
\item \textsuperscript{24} See e.g. Matthew L. Williams, Pete Burnap, and Luke Sloan, ‘Crime Sensing with Big Data: The Affordances and Limitations of Using Open Source Communications to Estimate Crime Patterns’, \textit{British Journal of Criminology}, 31 March 2016, doi: 10.1093/bjc/azw031, who found that the same Twitter data can reveal juxtaposed phenomena in areas with either high or low levels of crime.
\item \textsuperscript{25} Leslie, ‘Territoriality, Map-Mindedness, and the Politics of Place’, 172.
\end{itemize}
in order to bring a claim, for instance in a class action setting, claimants need to be aware that they belong to a group that has been negatively impacted. In the context of algorithmic grouping, where groups are constantly being made, re-made, and deleted on the basis of hypercomplex probabilities, this awareness is usually lacking.\(^{33}\)

Furthermore, relying on the current EU data protection framework often falls short. Consider for instance the idea of consent, a cornerstone in individual data protection law.\(^{34}\) It seems questionable whether the (often heavily intoxicated) visitors to the Stratumseind gave informed consent to use their personal data for research purposes. And what about their rights to receive access to the data kept on them, in order to see if it is correct,\(^{35}\) or whether so-called 'special categories' of data such as their ethnicity or sexual orientation - easily guessed by observing e.g. the entrance to an LGBT nightclub - have been processed?\(^{36}\) Both questions are, from a legal perspective, moot because the Living Lab and Strava anonymize all data before analysis. The latter even gives users an opt-out to making their anonymized data available.\(^{37}\) The individual focus of the human rights framework means that the analysis of anonymous data that by definition cannot be traced back to any individual can never infringe any data protection rights. However, it is clear that groups can suffer the consequences of the processing of data that does not identify anyone in the group, for instance when a map clearly indicates where extra police presence might be needed or where antagonists can strike secret military bases.

Many have argued that under these circumstances it no longer makes sense to only defend the personal data protection rights of individuals.\(^{38}\) Rather, we should be protecting group rights, which are rights ‘possessed by a group qua group rather than by its members severally’,\(^{39}\) and more specifically, the data protection rights of groups created by classification algorithms.

However, group rights are an awkward fit for the current European Data Protection Framework, which is heavily focused on individuals.\(^{40}\) Combined with the recent adoption of the GDPR

\(^{34}\) GDPR, point (a) of art 6(1).
\(^{35}\) ibid, art 15.
\(^{36}\) ibid, art 9.
\(^{40}\) Bart van der Sloot, ‘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 Linnet Taylor, Luciano Floridi, and Bart van der Sloot (eds), Group Privacy: New Challenges of Data Technologies,
and the associated legislative fatigue, which prevents any major innovations in data protection rights in the foreseeable future, it seems unlikely that group data protection rights will become a staple of data protection law in Europe anytime soon. This section has accentuated why this might be problematic. Therefore, the next section will introduce DPIAs, which might provide a workable solution that can be implemented without completely overhauling the legislative framework.

**Making Group Rights Real Through DPIAs?**

DPIAs have been around for several decades, but the GDPR has renewed interest in them. The GDPR became directly applicable in the member states of the European Union (EU) on 25 May 2018. It aims to create ‘first-rate data protection rules providing for the world’s highest standard of protection.’ Some of the central tools the GDPR employs in order to ensure this high standard are preventative measures, such as storage limitations, codes of conduct, certification, data protection by design and by default, and rules on the security of personal data. This approach aimed at reducing risks seems fitting for personal data protection, as it can be difficult to predict harms, and it might be even more complicated to reverse them.
DPIAs are a central instrument in this toolbox, and required if the processing of personal data is "likely to result in a high risk to the rights and freedoms of natural persons." In cases where it is unclear whether high risks will materialize, the Article 29 Working Party recommends that a DPIA be carried out nonetheless in order to minimize risks and ensure compliance. A DPIA is a process designed to describe the processing, assess its necessity and proportionality and help manage the risks to the rights and freedoms of natural persons resulting from the processing of personal data. This process can take many forms depending on, inter alia, the type of personal data processing being assessed. But its minimum requirements are a description of the processing; an assessment of the proportionality and necessity of the processing (i.e. can the same aim be achieved with less personal data processing); measures to minimize risks to data subjects; and an active involvement of those data subjects.

Crowdsourced maps are usually of publicly accessible areas such as streets, neighborhoods and parks, if only because it would be hard to find a crowd in an area that is not publicly accessible. In particular, the GDPR requires DPIAs if 'systematic monitoring of a publicly accessible area on a large scale' takes place. The Article 29 Working Party points out that such data collection can be especially problematic because data subjects might not be aware of the data processing. Furthermore, they might not be able to prevent their data from being collected without avoiding the public area in question, rendering the public place less public and any form of consent meaningless. Therefore, safeguards to gauge and minimize risks are certainly needed, and it seems it will be nigh on impossible to avoid doing a DPIA when gathering crowdsourced data for making maps.

As DPIAs are thus a necessary step in the development of crowdsourced maps, they might form a promising avenue to address the problems identified in the previous section. As we have seen however, trying to deal with group data protection rights within the EU data protection framework achieves unsatisfactory results. So why should Data Protection Impact Assessments be better suited to deal with mapmaking algorithms that analyze crowdsourced data if they are a part of that same framework? In the next two sub-sections, only those aspects that are pertinent to this specific question will be dealt with. Many others have already written extensively on DPIAs and proposed various models and frameworks. The object here is not

---

51 GDPR, art 35(1).
53 ibid., 31. Note that the Article 29 Working Party has been renamed the European Data Protection Board (EDPB) when the GDPR came into force. GDPR, art 68-76.
55 ibid, annex 2.
56 Although not completely impossible if one for instance invites a crowd of people to their private castle or estate.
57 GDPR, point (c) of art 35(3).
59 See, among many others, e.g. the interactive software released by the French Data Protection Authority, the Commission Nationale de l'Informatique et des Libertés, 'The Open Source PIA Software
to duplicate their work or add yet another model, but to suggest how, within existing models, a small extension could yield significant results.

**Opportunities**

Counter-intuitively, the embeddedness of DPIAs provides an opportunity to enhance the protection of group data protection rights within the current legal framework. Many other proposals exist to include various ethical, social, and human rights considerations in a plethora of impact assessment tools. These can serve as important inspirations and templates in the context of crowdsourced map-making initiatives. However, they form an additional financial and administrative burden that data controllers are called upon to voluntarily shoulder. By adding group rights to an already existing requirement, these costs could be significantly decreased.

The most straightforward way in which group data protection rights for crowdsourced map-making initiatives can be safeguarded is the same way in which personal data protection rights are safeguarded: by 'assessing [risks] and determining the measures to address them'.

Many already established personal data protection principles can be applied to address any identified risks for groups. Consider, for example, how the data minimization principle could also be applied to anonymous data. In the Strava example mentioned above this simple procedure could have prevented the company great reputational loss, not to mention the unknown costs to military operations.

This example also shows how important the proactive nature of DPIAs is in making group rights real. DPIAs should be engaged in 'as early as is practicable' in order to prevent risks from materializing. As was discussed in the section 'Maps and (Missing) Group Rights', it

---


62 ibid, 14.
can be difficult to reliably reconstruct what has happened once an algorithm has grouped individuals. Therefore, it is preferable to set limits and objectives beforehand, by for instance auditing algorithms for disparate impacts using simulated data.

Besides these more general opportunities for the inclusion of group data protection rights within DPIAs, the GDPR contains a clause with specific relevance for crowdsourced map-making initiatives: ‘where appropriate, the controller shall seek the views of data subjects or their representatives on the intended processing’. At its most basic, this clause allows data controllers to seek these views and perhaps include them in the final report of the DPIA, but then for all intents and purposes discount them. However, if properly engaged with, it also allows for data subjects to co-produce the environments in which they live.

This opportunity is created by the considerable performative power of maps, as discussed in the section ‘How Maps are Made’. By affording local groups access to the development process, they gain ownership over the production of the map and thus their surroundings. For instance, they can have a voice in what maps will and will not show, in which way, and to whom. This ability to contribute to the meaning of places makes DPIAs for crowdsourced maps especially well-suited for empowering residents; it allows them to change their environment from a place that they happen to live in and that others map, to a place that is mapped and co-produced by them.

Such a co-production can be modelled after consent, one of the pillars of personal data protection law. Consent should always be given beforehand, and by the affected data subject itself. Consent is only valid if, inter alia, it meets the connected requirements that it is informed, granular, and specific. By integrating these requirements into DPIAs, groups are given a way

63 Mittelstadt, ‘From Individual to Group Privacy in Big Data Analytics’, 489.
64 Goodman, ‘Discrimination, Data Sanitisation and Auditing in the GDPR’, 503.
65 GDPR, art 35(9).
67 To be understood in the manner that e.g. Michiel de Lange and Martijn de Waal use for the term; ‘Owning the City: New Media and Citizen Engagement in Urban Design’, First Monday 18.11 (27 November 2013), DOI: 10.5210/fm.v18i11.4954.
71 Although consent can also be withdrawn at any time (GDPR, art 7(3)); this can be difficult in this context as will be discussed below.
to meaningfully co-produce the crowdsourced maps and the GIS software - and thus their living environments - at a stage when some of the processing operations are still unknown.\textsuperscript{73}

The inclusion of as many views as possible is essential for this process. However, as it seems quite impractical to have all (potential) data subjects fully engaged in any and all DPIAs for crowdsourced mapmaking initiatives that might affect them, the selection of data subjects and their representatives is crucial. These should come from at least two categories: on the one hand representatives of local communities and local disadvantaged groups, and on the other hand, representatives from digital rights associations with a broader basis in society.

Delegates from local\textsuperscript{74} communities should be included to directly speak for those affected and in turn, affect the mapmaking process and the places they inhabit. Also, they alone can meaningfully and actively co-produce their environment on behalf of its inhabitants. The inclusion of representatives from disadvantaged groups can help in trying to avoid bias and discrimination. As was pointed out above, georeferenced data gathering tends to underrepresent already disadvantaged groups,\textsuperscript{75} and if data is collected on these groups, it usually further stigmatizes them.\textsuperscript{76} The inclusion of these groups should draw the attention of mapmakers to their specific concerns at an early stage.\textsuperscript{77}

The benefit of including digital rights associations is that they can represent groups of which the existence is not yet known and can contribute expert knowledge.\textsuperscript{78} As was pointed out in the section ‘Maps and (Missing) Group Rights’ it can be impossible to predict, or even reconstruct after the fact, which groups the algorithms constantly (re)generate. This would mean that any unforeseen groups are automatically excluded from the DPIA. In order to prevent this, digital rights associations may represent them. Note that it would seem appropriate for representatives of such associations to be cautious beyond this specific remit.

The selection of representatives is a difficult task,\textsuperscript{79} but the GDPR gives a hint at who might be welcome at the table regardless. Article 80 outlines how individual data subjects can mandate a ‘not-for-profit body, organisation or association which [...] has statutory objectives which are in the public interest, and is active in the field of the protection of data subjects’ rights and freedoms with regard to the protection of their personal data\textsuperscript{80} to represent them when lodging

\begin{footnotes}
\footnote{Article 29 Data Protection Working Party, ‘Guidelines on DPIA’, 14.}
\footnote{The exact scale of ‘local’ is not defined here, as it will depend on the mapmaking effort in question. When making a global map, ‘local’ might thus very well include the world population.}
\footnote{Pak, Chua and Moere, ‘FixMyStreet Brussels’.}
\footnote{Brannon, ‘Datafied and Divided’.}
\footnote{Wright and Mordini, ‘Privacy and Ethical Impact Assessment’, p. 402.}
\footnote{See also the next sub-section.}
\footnote{GDPR, art 80(1).}
\end{footnotes}
a complaint with a DPA, or when seeking a judicial remedy against a DPA, data controller, or processor. This list could be expanded - either informally by data controllers currently executing a DPIA, or eventually by the European legislator - to include the representation of data subjects during DPIAs in the sense of Article 35(9) of the GDPR.

**Limitations**

Although the EU legislator pays lip service to the ethical and social issues that result from large scale data processing and have an impact beyond the individual, these are included neither in Article 35 of the GDPR which describes Data Protection Impact Assessments, nor in the various DPIA models provided by the Data Protection Authorities (DPA) of the Member States. As a result, despite the many opportunities listed above, a number of important limitations to DPIAs as a tool for enhancing group data protection rights in the context of crowdsourced mapmaking needs to be considered.

A fundamental limitation to the possibility of using DPIAs as embedded in the GDPR is formed by the voluntary nature of the inclusion of groups and their representatives in the process. Article 35(9) is qualified as follows: *Where appropriate*, the controller shall seek the views of data subjects or their representatives on the intended processing […]. It remains unclear, at least for the moment, where this would and would not be appropriate; a situation that the Article 29 Working Party failed to remedy in its opinion on DPIAs. It seems improbable that companies will interpret this provision widely and be eager to engage in the time-consuming, costly, and potentially politically laden process if it can easily be avoided. As seen above, however, it would be exactly this inclusion that engenders many opportunities.

If we assume that companies do engage in DPIAs and include affected data subjects and representatives of local (disadvantaged) groups, many limitations still remain. First of all, when we compare the data protection rights of groups that can be protected through DPIAs to the rights that individuals have over their personal data, it seems that it is chiefly the prohibition on the processing of special categories of data that can be somewhat enhanced through engaging in DPIAs. Left by the wayside are many other principles, such as for example accuracy, accountability, confidentiality, or a lawful basis for processing such as informed and freely given consent. For now, these data protection principles seem out of reach for groups.

---

81 ibid, art 77.
82 ibid, art 78 and 79.
83 Mantelero, ‘AI and Big Data’, 756; GDPR, recital 75. See also footnote 61.
85 Mantelero, ‘AI and Big Data’, 755.
86 See GDPR, art 5 for a more comprehensive set of data processing principles.
The selection of groups to be represented would also present a data controller eager to conduct a DPIA with major difficulties. As was discussed in the section ‘How Maps are Made’ it is impossible to pinpoint exactly who produces, edits, or consumes a contemporary map; we could add to this confusion those who are impacted by a map. Even if it would be possible to neatly separate these roles and find suitable representatives for each affected group, the composition of all the groups that will (or might) be formed by an algorithm cannot always be known beforehand. Furthermore, many distinct local communities are not neatly divided in classical neighborhoods, and the involvement of neighborhood associations, which may be easy to access for data controllers, might not lead to accurate representation. Finally, assuming that a somewhat complete overview of groups to be represented has been established, it is extremely difficult to decide who can speak on behalf of each group. As the success of a DPIA hinges on the accurate composition of the groups and involvement of their representatives, careful consideration for each separate DPIA is warranted.

Finally, even if companies do engage in the process initially and include as many views as possible, for a DPIA to be truly successful it should be a circular process that is regularly repeated. This is even more important if self-learning algorithms are employed as their outcomes can show unexpected changes over time. This can include changes in the groups targeted, thus necessitating a constant updating of the composition of representatives. The costs necessary to continually assess the possible negative impacts of crowdsourced map-making software that has already been written and released might not, in the view of profit maximizing companies, be justified by the possible results.

Conclusion

The question asked at the beginning of the chapter was whether DPIAs could form a panacea or a Band-Aid when protecting group data protection rights in crowdsourced mapmaking initiatives. It laid out the strong performative power of maps and how crowdsourced data is used to make them. Then, it introduced how this practice interacts with the current personal data protection framework in the European Union, leaving undesirable gaps in the safeguarding of group rights. Finally, it introduced Data Protection Impact Assessments and discussed how they could help in alleviating these problems without overhauling the current legislative framework in the EU.

When taking stock of both the opportunities and limitations that DPIAs offer for group data protection rights in crowdsourced mapmaking initiatives it seems that they could easily be circumvented, their effect would be limited at best, that the proper representation of groups is nigh impossible, and that their long-term impact is uncertain in the face of self-learning.

89 The permanent ad-hoc committees proposed by Mantelero might provide some solace for specific data controllers, although it remains unclear who would be responsible for solving the underlying problem: ‘AI and Big Data’, 771.
algorithms. Still, companies interested in retaining consumer trust - and gaining a competitive advantage when dealing with responsible customers and partners - would be well-advised to make the investment. This goes doubly so for governmental bodies and institutions: DPIAs form an opportunity to use group rights to put the performative power of maps in the hands of those being mapped. Despite the many gaps and pitfalls, DPIAs for mapmaking initiatives that utilize crowdsourced georeferenced data should be performed, and they should include as many views as possible; Public space should belong to the public, not to companies writing mapping software.

It is up to the EU legislator - and in the meantime: the European Data Protection Board, formerly known as the Article 29 Working Party, whose importance ‘for the EU data protection cannot be overstated’\textsuperscript{90} - to ensure that group data protection rights can be properly incorporated within the European Data Protection framework. This is the only way to ensure that the uncomfortable Band-Aid proposed in this chapter becomes unnecessary and can be ripped off.

Acknowledgments

An early version of this paper was presented to my colleagues at the department of Trans-boundary Legal Studies in Groningen. I thank them for the frank comments, which helped to sharpen the argument. I thank Prof. Jeanne Mifsud Bonnici for her many comments on various early drafts. Thanks are due to Lauren Elrick for her careful language check. Finally, I would like to thank the participants of the Surveillance Studies Network 8\textsuperscript{th} Biennial Conference at Aarhus University, Denmark, who provided valuable feedback and inspiration.

References


Mittelstadt, Brent. 'From Individual to Group Privacy in Big Data Analytics', *Philosophy & Technology* 30.4 (December 2017): 475-94, DOI: 10.1007/s13347-017-0253-7.


OpenStreetMap Foundation. 'About', https://www.openstreetmap.org/about.

Pak, Burak, Alvin Chua, and Andrew Vande Moere. 'FixMyStreet Brussels: Socio-Demographic Inequality in Crowdsourced Civic Participation', *Journal of Urban Technology* (10 April 2017), 1-23, DOI: 10.1080/10630732.2016.1270047.


