How the GDPR can contribute to improving geographical research

Louise Meijering, Tess Osborne, Esther Hoorn, Cristina Montagner

1. Introduction: The GDPR and research

The General Data Protection Regulation (GDPR) took effect in May 2018 and provides data subjects improved rights and political agency over their personal data. Personal data are data that can be used to identify a natural person. The introduction of the GDPR was a response to the plethora of personal data that are collected, used, and often sold by multinational companies, such as Facebook and Google, without effective transparency and understandable terms and conditions, as well as to data breach scandals (e.g. Cambridge Analytica) (see Kitchin, 2014; Paganoni, 2019; Politou et al., 2018). Not only does the GDPR improve individual agency over one's data, but it also plays an important role in harmonizing the data protection laws across the EU as it de-emphasises nationality and residency (Clark and Jones, 2008; Paganoni, 2019).

The GDPR could be considered sympathetic to research processes since it acknowledges that data produced from research has validity temporally and across projects (Rec. 156). However, many sector-related methodological and ethical standards are still under construction, and as a result the extent of exceptions for research purposes remains unclear in a variety of situations (Tene and Polonetsky, 2016). Moreover, the harmonization intended by the GDPR across the EU is limited for scientific research because to some extent it is left to the member states to implement exceptions (Art. 9(2)(j)). From a researcher’s perspective, negative potential impacts of the GDPR include "regulatory ambiguities […], their compatibility with research requirements for obtaining consent and additional research burden for researchers and research institutions" (Health Ethics and Policy Lab, 2019: 60). These have led to an increasing number of scholars expressing their concern and angst towards the GDPR, especially in the medical and data sciences (Cornock, 2018; McCall, 2018; Staunton et al. 2019). Despite these ongoing discussions, there is yet to be a reflection on how the GDPR affects research from a geographical perspective. Aside from contributing to the critical discussion on the GDPR, the paper also builds upon the debates within geography on the use of big data. Such debates centre around how big data contributes to reproduce and reinforce inequalities, for instance in biogeography, smart farming, and the digital representation of space (Ebach et al., 2016; Lambio and Lakes, 2017; Lioutas and Charatsari, 2020). Also, the ethical, epistemological, and methodological questions raised when working with this type of data are subject of discussion (Kitchin, 2013).

With this Forum paper, we aim to (1) discuss our experiences of how the GDPR affects research from a geographical perspective; and, in doing so, to (2) suggest good practices for conducting research in the era of the GDPR, specific for the field of geography. We reflect upon the privacy implications of two research projects that we have begun since the GDPR took effect. Our projects have three characteristics that make them typical in the field of geography. First, they...
A DPIA is a self-assessment exercise to identify and reduce the risks associated with the use, management, and processing of data (Binns, 2017). In research projects that collect and process personal data, a DPIA is often mandatory. According to the GDPR, the purpose of a DPIA is “to identify possible data breaches” and “high-risk data processing”, assess the risks related to the rights of data subject, and come with mitigating measures (Art. 35). In the context of a research project, a DPIA is a collaborative effort of researchers, data consultants, legal and security advisors, data subjects, and privacy coordinators, and is led by a Principal Investigator (PI).

Several models that serve to guide the DPIA process have been developed, although none of these have been specifically designed for research projects (see, for example, Bieker et al., 2016; Ministerie van BZK, 2017). While conducting the DPIAs for our research projects, we followed the model developed by Bieker et al. (2016), which is based on six privacy and security protection goals: confidentiality, integrity, availability, unlinkability, intervenability, and transparency. A DPIA assesses the risks associated with these goals and to produce measures to mitigate these risks. As such, a DPIA enables researchers to effectively embed privacy in the project’s design and throughout its life cycle.

have an international character, where we conduct fieldwork and collaborate with researchers in different countries. Second, our research design includes a mixed-methods approach, where we combine spatial (GPS) and qualitative data. Third, we typically work with participants who can be characterised as ‘vulnerable’.

In assessing the privacy implications of two of our externally funded research projects we conducted a Data Protection Impact Assessment (DPIA). A DPIA aims to identify privacy risks in a (research) project, and to develop measures to mitigate these risks (Box 1). While conducting the DPIAs we identified three issues that can have a significant impact on geographical research; namely, informing participants, data management, and international collaboration. We explore each issue in turn, reflecting upon the various techniques we adopted to mitigate the impact they had on our research planning and conduct. To conclude, we discuss the benefits and difficulties that arise from our approach to data protection and present discipline-specific good practices.

2. Informing research participants

The GDPR requires that a variety of information should be provided to research participants, including the (lawful) purposes for processing the personal data; the storage period for that data; and how the data will be shared and stored; and the contact details for data processors (ICO, 2019). However, there is an inevitable challenge implied in having to provide a vast amount of complex information in a digestible way.

In our projects, where we work with older participants who experience memory loss and/or have suffered a stroke, we had to develop ‘GDPR-proof’ information and consent forms. This resulted in a first draft that consisted of eleven-pages of text and despite our sincere attempt to give thorough, detailed, and transparent information to the participants, the information was perceived as overwhelming by our potential participants. While participants may be fully involved and informed (legally) with ‘GDPR-proof’ information, they may understand less of what they are consenting to and/or what the research is about. Specifically, for people with dementia or memory problems, it has been argued that they are more vulnerable to “information overload” meaning that obtaining informed consent can be particularly difficult (Novitzky et al., 2019: 178). In our projects, we reduced the information as much as possible but without compromising the GDPR requirements with two main strategies: (1) providing information in different layers; and (2) supporting text with infographics to facilitate understanding (Fig. 1).

In one of our projects, we developed four distinct layers of information that potential participants are provided with. The first layer is a one-page leaflet with the announcement that we are looking for research participants, information on inclusion criteria and what participation entails, supported with infographics and a picture, and contact information including logos. The second layer is a two-pager, which contains a bit more written information about the project as well as photos of the research team. The third layer consists of an invitation letter with information sheet and informed consent form. These are again supported with the visual material and the infographics that were used in the first two layers of information. The information sheet contains information about the project and its aims, participation and data collection, risks of participation, what we do with the participants’ data, participants’ rights and contact information. The fourth layer consists of more technical information on data management and privacy, which is referred to in the information sheets and available through the project’s website. This is information that participants are typically hardly interested in or find difficult to take in. An important reflection is that we see our printed material as having a supportive role, and that it complements the verbal information and explanations that we give to (potential) participants.

3. Data management

Additionally, we identified risks connected to the management of data collected from our research participants. The data management conditions of the GDPR align with the rulings of research funders and universities, who require data management assessments and agreements to be made before commencement of data collection in a Data Management Plan (DMP). A DMP aligns with the DPIA measures. It details what data are collected and from whom; as well as how and where the data are transferred, stored, and exchanged for the duration and aftermath of the project (Michener, 2015).

Creating and maintaining a DMP is useful, but we believe that the research community could benefit from additional guidance as to what constitutes a good DMP. In this context, it is worthwhile to mention that, to-date, we have not received feedback on any DMP we have written. This is a missed opportunity, as we are convinced the quality of our DMPs would be enhanced by feedback from other people beyond our collaborators and advisers. We would therefore recommend to streamline the DMP process and developing a thorough system through which feedback is given and incorporated.

An important part of data management is being able to store and work with data in a secure work environment. To accommodate this, our ICT department developed a so-called Virtual Research Workspace (VRW). The VRW enables us to store and analyse personal data in a secure way. It requires two-factor authentication (password and text-message). Also, the internet cannot be accessed when working on it. Additionally, people with access to the VRW are allocated different rights associated with reading or writing files, as well as up- and downloading rights (Table 1). Finally, access rights to different folders and files can be managed by the PI.

While we are lucky to have access to such a workspace, we are one of the first groups to work on the VRW and have encountered some of the ‘teething troubles’ with a new system. For instance, we encounter problems with uploading documents, accessing other drives, and getting the software we need (such as ArcGIS) onto the VRW. Additionally, we are charged usage fees for using the VRW; and while this is paid through our project budgets, not all researchers have the luxury of working in externally funded projects, and therefore not the means to pay for a VRW – even if the GDPR might require this.
4. International collaboration

The field of geography prides itself on collecting data and working with collaborators around the world. Consequently, the transfer of data across national borders is common in geographical research. As we discussed in the introduction, the GDPR allows member states to implement exceptions. This may lead to inter-country differences in the law, which in turn make international fieldwork and academic collaboration more difficult (EPRS, 2019). In the case of collaborations and data collection beyond the EU, it is important to note that GDPR safeguards apply to all personal data that are processed in the EU, and to all personal data collected from European citizens, irrespective of the place where they are processed (Art. 3). This means that, if personal data are transferred either into or out of the EU, the legislative power of GDPR applies. This may sound straightforward but we found that this is sometimes difficult to put in practice.

With regard to data collection beyond the EU, data transfer back
into the EU can be tricky. For example, India’s Data Protection Bill imposes restrictions on exporting personal data, means that only pseudonymised data can leave the country. Additionally, and following Brexit, the UK has forbidden the import of personal data into the country (personal communication with Lancaster legal officer). Such restrictions pose challenges to international collaborations, where we work with personal data that are difficult to pseudonymise, such as location data.

Another challenge we encountered while conducting fieldwork beyond the EU, is that as EU-based researchers, we are obliged to provide our research participants both inside and beyond the EU the detailed information discussed in Section 2. As participants outside the EU are not accustomed to the GDPR as the legislative context, our ‘GDPR-proof’ information sheets and consent forms may be even more overwhelming than for EU-based participants, and thus weaken the trust and rapport between researcher and researched.

Furthermore, we have encountered different understandings of data protection, ethics and privacy during our collaborations with researchers beyond the EU (where the GDPR does not apply). For instance, for one of our collaborators, ethical issues centred around inclusion criteria for participants, whereas we were more focused on informed consent. After this realisation, we have sought ways to overcome these differences. In so doing, we decided to start from our joint interest: one of our research objectives, which is to compare the housing experiences and opportunities for older adults who experience memory issues in three different welfare contexts. We organised a session, to discuss how to go about exchanging and jointly analysing data to achieve this objective. Our discussion was informed by the legal and ethical issues and possible measures that we had identified in the DPIA process, and helped us to consider how to achieve international comparison in the best possible way. Although we recognise the value of guidance on international data transfer (EDPB, 2018), discussions between academics ‘on the ground’ are vital in making data exchange a success. Overall, we feel that different legal and ethical frameworks do make international collaboration more difficult, because they result in considerable time and effort that needs to be spent on harmonising these frameworks in research practice.

5. Discussion and conclusion

Through our discussion around the Data Protection Impact Assessment (DPIA) process, we have discussed how informing research participants, data management, and international collaboration can be impacted by the GDPR. By reflecting on our experiences of the GDPR and the planning and conduct of our geographical research we have devised good practices specific for the discipline, and enriched the discussion on the GDPR’s influence on research in general.

With regard to informing research participants, we found that the GDPR encourages researchers to inform their participants thoroughly, and provide participants with more agency over their data (e.g. to enable them, for instance, to withdraw whenever they want). This should contribute to building trust and rapport. However, we found that participants can feel overwhelmed when presented with such thorough information leading to mistrust and a lack of autonomy (see also Health Ethics and Policy Lab, 2019). We tried to resolve this issue by presenting participants with layered information, supported by infographics and gradually building up in complexity. This ties in with other efforts, such as around the use of animations in informed consent (Hong et al., 2012). Then, in our experiences around data management, including transfer, storage, analysis and sharing, we found the Data Management Plan (DMP) a potentially useful tool that could be further enhanced by expert feedback. Having a secure workspace for data storage and analysis is vital for the safety of data, but is also fraught with practical challenges, for instance around software availability. Finally, turning to international fieldwork, we discussed the implications of working with participants and collecting data outside the EU as well as the associated data transfer issues. Furthermore, we have demonstrated how the different ethical and legal frameworks within and beyond the EU can hinder effective collaboration. Our advice would be to start from the researchers’ perspectives and needs in addressing this.

As is common throughout the academic community, we are supported in addressing privacy-issues by our Research Data Office (RDO) and legal advisors. Such support for researchers is now fundamental, not only for good research practice but also because funding agencies, such as the ERC demand a DPIA in the data management and ethical requirements for projects (JAPP, 2019). This underlines the fact that addressing privacy issues in academic research is a shared responsibility between researchers and their institutions: as institutions require researchers to acquire external research funding, they have a duty of care to make running such projects possible. In line with this, the university made e-learning material available that enables practice with the principles of the GDPR (University of Groningen, 2020).

A general issue that we want to raise, is the time-investment that is needed to address the privacy issues relevant to (geographical) research projects. In our experience, this time-investment is high: as researchers, we have to gain understanding of relevant privacy issues and ways to tackle these. Furthermore, we have to explain the particulars of our research projects to support staff to enable them to give effective advice. Although we do see the benefits of addressing the privacy implications of research projects at an early stage, the workload associated with it is high and not always manageable. We would like to position this issue more broadly in debates around the neoliberalist university in geography (Berg et al., 2016; Riding et al, 2019). In this context, addressing privacy issues in research is often yet another task that needs to be done in everyday realities of competition and production and adds to the ever-increasing pressure placed on academic staff.

Whilst we are critical of the DPIA process, we do acknowledge the strengths of the process. Data security, clarity with regard to processor responsibilities, and research collaborations within the EU are enhanced, as well as the autonomy of research participants. As the DPIA is a new method in geographical research, it is a learning process that requires a significant amount of work. To reduce the effort and the workload for future researchers and support staff, we suggest that is it important to establish ethical codes of conduct and standard procedures with so-called reference DPIAs for research scenarios. We hope to have contributed an initial building block towards this to demonstrate that the GDPR does not make geographical research impossible, but rather contributes to making it more effective and fair.

Funding statement

Louise Meijering and Tess Osborne would like to acknowledge the funding they have received from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Meaningful Mobility, Grant Agreement No. 802202). Louise Meijering also acknowledges the funding she has received from the Joint Programming Initiative More Years, Better Lives, carried out by the Netherlands Organisation for Health Research and Development (COORDINATES, Project No. 9003037412).

References

Cormock, M., 2018. General Data Protection Regulation (GDPR) and implications for