ISSN 2704-8012

European Journal of Privacy Law & Technologies

2019/2

G. Giappichelli Editore

European Journal of Privacy Law & Technologies

Directed by Lucilla Gatt

2019/2

G. Giappichelli Editore

European Journal of Privacy Law & Technologies
On line journal
Italian R.O.C. n. 25223

G. GIAPPICHELLI EDITORE - TORINO
VIA PO, 21 - TEL. 011-81.53.111 - FAX 011-81.25.100
http://www.giappichelli.it

The Journal is one of the results of the European project TAtoDPR (Training Activities to Implement the Data Protection Reform) that has received funding from the European Union's within the REC (Rights, Equality and Citizenship) Programme, under Grant Agreement No. 769191.

Published Online by G. Giappichelli in Dicember 2019
www.ejplt.tatodpr.eu

DOES PRIVACY BY DEFAULT MEAN RESEARCHERS SHOULD RECONSIDER RESEARCH ETHICS PRACTICE IN RELATION TO RECORDING INFORMED CONSENT

Alex Nunn

Abstract

It is normal practice for researchers collecting data from 'human subjects' to record that their participants have provided 'informed consent'. This often means recording personal data such as a name, address and signature when the underpinning research question – or in legal terms 'the specific purpose for processing data' does not actually require this. The provisions of GDPR in relation to 'privacy by default' might provide a rationale to revisit normal practice and ethical guidelines to give greater emphasis to anonymisation or pseudonymisation at the point of data collection. It is recommended that research organisations and researchers revisit normal practices and guidelines to consider where anonymised data collection might be utilised more fully.

1. GDPR a Summary

From May 2018, Europe's new data protection regime was significantly tightened with the introduction of the General Data Protection Regulation (GDPR). The impact of this regulation cannot be overstated – it touches all areas of organisational functions and substantially strengthens the rights of data subjects in relation to data controllers and processors. Proponents of GDPR suggest that for those already practicing strong data protection, the changes introduced by the GDPR are incremental. But even where incremental, some important challenges are thrown up by the changes.

The 99 articles of GDPR establish 8 rights for individuals: (1) to be informed, (2) access to the data held about them, (3) to have errors corrected, (4) to have their data erased, (5) to restrict aspects of data processing, (6) to be able to

move their data to a different organisation, (7) to object to the use of their data and (8) in relation to automated data profiling. GDPR ensures that organisations must have an appropriate 'legal basis' for collecting, storing and processing individuals' personal data and establishes the different possible grounds on which this may be constructed. For most organisations, these include consent, the delivery of a contractual services or to uphold the law in other respects or to protect 'vital interests' such as life preservation. [1]

The issue of consent is very familiar to researchers involved in collecting data. As with most extant guides and good practice in research ethics, consent in relation to GDPR must satisfy the criteria that it is given freely, is informed, specific and explicit. It must also relate to the collection, storage and use of the data, including any transfer of data between organisations. There are also specific measures related to children, with 'children' defined as those under the age of 18.

2. Research Ethics – The Position before the GDPR

Many of the principles of the GDPR are not new to researchers. As discussed below, the principles of informed consent, offering opportunities for withdrawal and early anonymisation of data are all well entrenched in good practice guidelines. For researchers, data protection usually arises in the context of 'research ethics'.

It is normal research practice in research involving human participants for details of the research design to be put in front of an ethics committee and be approved by that committee prior to any data being collected. For example, the UK Research Integrity Office suggests that

> *Researchers should submit research projects involving human participants, human material or personal data for review by all relevant ethics committees and abide by the outcome of those reviews. They should also ensure that such research projects have been approved by all applicable bodies, ethical, regulatory or otherwise.* [Section 3.7.9.] [2]

[1] Information Commissioner's Office, 'Guide to the General Data Protection Regulation (GDPR)' (2019) <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/> accessed 23 October 2019.

[2] UK Research Integrity Office, 'Code of Practice for Research' (2019) <https://ukrio.org/publications/code-of-practice-for-research/3-0-standards-for-organisations-and-researchers/3-7-research involving-human-participants-human-material-or-personal-data/> accessed 23 October 2019.

It is also normal that such ethics committees will require a sense of what questions the research will address, who will be involved and in what circumstances the research will be conducted. In practice, this involves submitting a summary of the research for approval, draft interview schedules or questionnaires, a sampling strategy and a plan for the ways that data will be collected including locations and timeframes. Prior to the GDPR, this usually included an 'informed consent' sheet to be given to research participants making them aware of the use of their data and how to withdraw from the research. In research projects where participants are at risk of some immediate harm – such as in clinical trials, these standards are particularly tight, but they are also applied to research investigating social or political issues. For example, the Social Research Association's Ethical Guidelines, currently dating from 2003 but due for review in 2019, state that informed consent "should ideally be both ORALLY and in WRITING", [3] emphasis in the original]. While many guidelines stop short of determining how consent should be recorded – this too has tended to be in writing, though most ethics guides stop short of mandating this. For example, the Economic and Social Research Council online Research Ethics Framework includes a 'Frequently Asked Question': "Is written consent always necessary?" which includes the following answer:

> *It is sometimes argued that formal written consent is not necessary because by consenting to see the researcher, a participant is in fact giving consent. However, it is good practice where possible for all participants to be provided with information giving the name and status of the researcher carrying out the study, a brief rationale of the study (including its purpose and value), and an account of why the individual is being invited to take part.*
>
> *The person interviewed should be made aware what will happen to the data, whether and how it may be shared with others, and whether they will be identified – and asked their preference.* [4]

At the same time, these various guidelines also imply that the responsibility to be able to demonstrate that consent has been gained belongs to the researcher. As such, despite this ambiguity in sectoral guidelines, most University ethical guidelines tend toward the default position that written consent is the expected norm, and that this is particularly the case where research involves risk and po-

[3] Social Research Association, 'Ethical Guidelines 2003' (2003) <https://the-sra.org.uk/common/Uploaded%20files/ethical%20guidelines%202003.pdf> accessed 23 October 2019.

[4] Economic and Social Research Council, 'Is Written Consent Always Necessary? - Research Ethics Framework Website' (2019) <https://esrc.ukri.org/funding/guidance-for-applicants/research-ethics/frequently-raised-questions/is-written-consent-always-necessary/> accessed 23 October 2019.

tential for harm, for instance by focussing on 'vulnerable' groups, such as young people, disabled people, or those experiencing social exclusion. Since social research frequently does focus on these groups, this is a common experience.

The default expectation then when researchers seek ethical approval for their data collection is that they will provide research participants with written information about the reasons for collecting data, how it will be stored, what analysis will be applied to it, how it will be placed in the public domain (e.g. confidentiality, anonymisation and pseudonymisation) and details about how they can withdraw. The right to withdraw in particular is often stressed as a condition for making abstract commitments to ensuring that consent is not a one-off process but an ongoing one, that may be subject to renegotiation. The Social Policy Association's guidelines are typical:

> *Consent to participate in a research study should be regarded as an on-going process and it should be made clear to participants that they are free to withdraw from the study or withhold information at any point. Participants should be given the opportunity to ask for further information about the study at any time.* [5]

This default position is also that researchers will secure written consent from their research participants. Here the University of Manchester's online guidance to staff researchers is illustrative. It suggests that consent can be gained in written or oral form but if it is the latter

> *Provided by asking the participant a series of questions (through the use of a consent script) and recording their verbal agreement to each statement. The recording can by done either by audio recording or through the use of detailed fieldnotes. If fieldnotes are used, you must include the participant's name, the date in which consent is being taken and the specific statements they are agreeing to. Please also note that if using this method you must provide justification to the ethics committee why this is needed.* [6]

What stands out here is both that not requiring written consent is to be regarded as an exception from the norm to be specially justified and that even where this is the case written notes of a participant name is required.

---

[5] Social Policy Association, 'SPA Guidelines on Research Ethics' (2019) <http://www.socialpolicy.org.uk/downloads/SPA_code_ethics_jan09.pdf> accessed 23 October 2019.

[6] University of Manchester, 'Preparing an Ethics Application' (2019) <https://www.staffnet.manchester.ac.uk/rbe/ethics-integrity/ethics/app-prep/#> accessed 23 October 2019.

3. Impact of Data Protection on Research Ethics Procedures

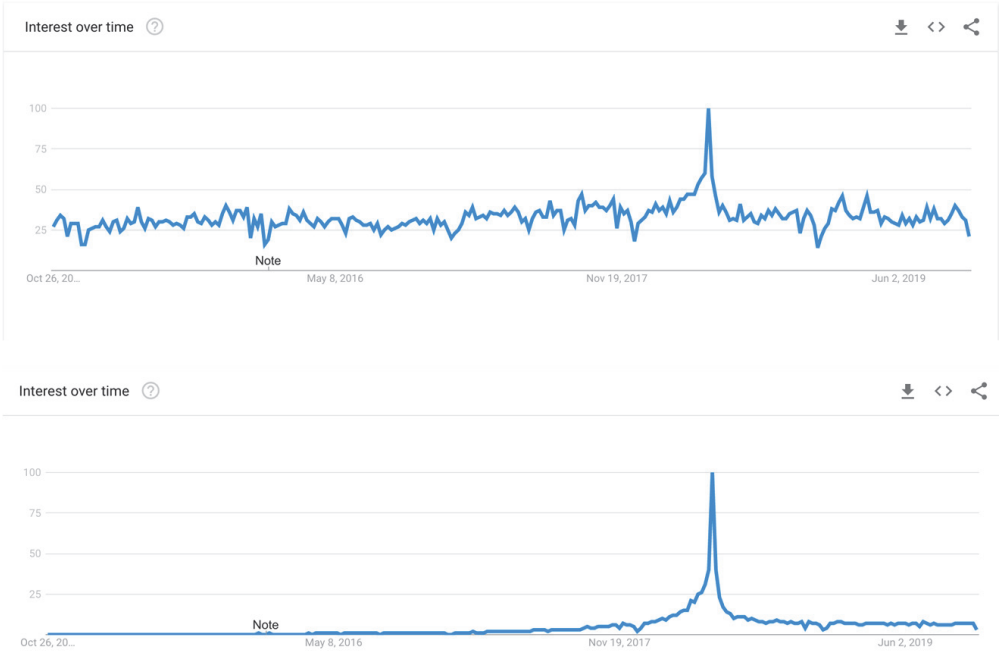
The introduction of the GDPR has certainly focussed minds in terms of data protection. Use of Google search data shows the frequency of searches for both Data Protection and GDPR specifically spiked around the time of the introduction of the regulation and most organisations will have undertaken training of their staff about GDPR, with many new organisational rules introduced as a consequence. Many of us will be familiar with the phenomena of organisational cultures of anxiety regarding data protection in the wake of the introduction of the GDPR, often leading to the imposition of rather overstated restrictions. However, in truth, in relation to the treatment of personal data in research projects, the requirements of pre-existing legislation (e.g. the Data Protection Act 1998) contained many of the same provisions.

However, the GDPR does enhance these requirements, particularly via what is often referred to as 'privacy by design' or 'default'. Article 25(1) of GDPR suggests:

> *Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.'*

The area of concern here is that of the treatment of personal data. Wherever a researcher has access to, or collects personal data then the provisions of GDPR and the domestic legislation associated with it (in the case of the UK the Data Protection Act 2018) applies – an appropriate legal basis is required and protections should apply to the collection, storage and processing of data. The GDPR also suggests that personal data collection should be minimised.

Figure 1: *Google searches last five years: Data Protection and General Data Protection Regulation*



Source: Google Trends, accessed 23-10-2019

4. Social Research and Personal Data

Most often, in social research, researchers are relatively uninterested in personal data as a component of their analysis. They may be interested in analysing their data according to criteria that might be used to construct personal profiles at the level of the individual such as ethnicity, gender, disability, area of residence and so on, but usually these are as abstract categories rather than as belonging to that specific individual. The specific individual is present in the research as some kind of 'representative' (even if the quantitative requirements for statistical representativeness is invoked in the research method applied) of the broader social group. The focus is on whether or not differences of gender, ethnicity and so on, influence aspects of the analysis rather than on the specific individual.

The exceptions here are two-fold. First where the research follows a longitudinal design to track changes over time with the same participants, usually where a specific group are subject to some form of intervention or where they are involved in large-scale cohort studies. However, longitudinal designs are

relatively rare in social research because they are administratively challenging and expensive to undertake. In most cases, even where change over time is important, this is explored *via* cross-sectional designs where different samples with similar abstract characteristics are used at different points in time. While in longitudinal designs the specific individuals are important, this is not the case in cross-sectional research.

The second exception is in fact accidental; usually where a confluence of abstract categories means that specific individuals are identifiable. This is the case for example when isolating locational (e.g. small areas of residence) and identity (e.g. gender, age or ethnicity). Recognising that area of residence is a strong factor or predictor of deprivation, small area identifiers (e.g. in the UK Super Output Areas derived from post-codes) and the overlap of these with other factors of deprivation such as age or ethnicity means that even though the focus of research may be on abstract categories which predict or result from deprivation, an accidental implication of this is that recognisable individuals are identifiable in the data.

All that said, for the most part researchers are not interested in personal data for their analysis. In most studies research participants – or data subjects – might be anonymised at the point of data collection, at least in regard to the analytical purpose of the research. However, in the main, because of deeply engrained ethical practices, researchers often unintentionally collect personal data as a product of demonstrating consent and ensuring the right to withdraw. In simple terms; there is a trade-off between administrative requirements to demonstrate ethical practice and data protection requirements to ensure that personal data is treated in line with the requirements of the GDPR (and indeed the predecessor legislation). Frequently researchers may be perfectly satisfied with collecting anonymised or pseudonymised data at the point of collection, but unwittingly turn this into personal data because of the requirement to stay within research ethics guidelines – or, more accurately – normal practice, because as we have seen sectoral guidelines and University procedures often stop short of formally requiring written consent.

The foregoing practices come with considerable costs in terms of administration, resources and the data infrastructure required. For example, because personal data has been collected, researchers are under an obligation to undertake elaborate administrative procedures such as replacing names and identifying information in their data with 'Unique Identifiers' and maintaining a separate record of how these Unique Identifiers relate back to specific individuals, so that promises of the right to withdraw can be upheld. In turn, this means storing data on separate parts of data management systems, and using encrypted data recorders and storage devices while 'in field'. They also come with risks attached.

What we know about human behaviour is that individuals with limited time, facing different pressures and the need to prioritise tasks often take shortcuts. The uncomfortable reality is that many researchers will routinely take risks of non-compliance, taking time to apply unique identifiers, forgetting to separate names and other identifying information or carrying data on unprotected or unencrypted hard drives, laptops and data recorders. Recognising the reality of these human frailties in data protection systems is an important step in minimising risks.

However, here GDPR actually provides a means of resolving the tension between engrained research ethics practice and legal data protection requirements, and the associated costs and risks. According to Article 25(2), ensuring 'privacy by default' means:

> *appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility.*

5. Conclusion

If data is being collected for the general purpose of research, but the research question being addressed does not require individual personal data, then the 'specific purpose of the processing' in turn does not require personal data. As such, it may well be opportune to use the new world of the GDPR requirements – and the enhanced obligation to seek 'privacy by default' - as a trigger to rethink research ethics. Specifically, universities and sectoral organisations might want to consider whether to place a greater emphasis on alternative methods of gaining consent which do not inadvertently turn research data into 'personal data'.

The essential question to ask is: 'can the data collected be anonymised at the point of collection, so that no personal data is ever stored?'. Clearly there are further trade-offs here. The right to withdraw after the point of data collection is somewhat compromised here – but providing participants with a unique identifier on their 'information sheet' and only recording this with the data collected in the first place provides a simple mechanism to maintain this.

It also introduces a risk that researchers may not seek consent, or that administrative structures do not trust researchers when recording consent without a 'signature' from a research participant. But here Universities and other research organisations need to consider whether it is more likely that researchers will not

comply with the relatively easy step of ensuring consent – a practice that is deeply engrained in behavioural norms – or that they may take short-cuts in time-consuming data protection practices where they do collect personal data. For the most part, it may well be that unsigned but uniquely numbered unsigned informed consent forms without any record of personal data be sufficient for ensuring administrative requirements while minimising personal data collection from the outset.