European Parliament

How the General Data Protection Regulation changes the rules for scientific research

The EU General Data Protection Regulation (GDPR) has the potential to facilitate and improve scientific research. However, in order realise this potential, policy options must be considered in three areas: regulations, procedures and practice; and there is also a need to address capacity building.

Regulatory options	Reconcile the requirement for specific, informed and free consent in the GDPR with the need for broad consent in scientific research and reconcile the definition of specific, informed and free consent with consent requirement in associated instruments.
	Clarify the derogations under Article 89(1) with respect to permitted processing for statistical and scientific purposes.
	Establish guidelines for the use and monitoring of anonymisation and pseudonymisation technology within different contexts.
	Develop institutional guidelines to assist researchers involved in transnational transfers of personal data outside the EU in the context of transnational collaborative scientific projects.
	Resolve the conflict between data subject rights under the GDPR (the right to data portability and the right to access data), and the protection of database rights under the *sui generis* regime.
	Coordinate the monitoring of derogations that apply to research and develop codes of conduct that address harmonisation deficits.
	Develop consistent standards of correctness and accuracy for all domains of scientific research with respect to data processing principles.
Procedural options	Install robust data management practices.
	Develop adaptive data governance frameworks.
	Develop technical standards for anonymisation and pseudonymisation based on best practice.
	Develop researcher-friendly software tools for GDPR compliance with a special focus on open access tools for data portability.
Transitional capacity-building options	Organise educational activities and specific training sessions for data protection literacy among researchers, students and scientific trainees.
	Support more research into post hoc impact assessment.
	Monitor attitudes and tailored data protection literacy interventions.
	Raise awareness about GDPR rights and obligations by means of activities for the general public.

EPRS | European Parliamentary Research Service

Scientific Foresight Unit (STOA)

PE 634.447 – July 2019

EN

1. Regulatory options

In order to reconcile the GDPR requirement for specific, informed and free consent with the need for broad consent in scientific research, one potential regulatory policy option would involve amending Recital 33 so that it explicitly acknowledges alternatives to specific informed consent. Another policy option would be to collect dynamic consent from data subjects. Although the GDPR requires informed consent for some forms of scientific research project, it may be impossible to determine how the data collected is to be used. Recital 33 of the GDPR notes that data subjects should be able to consent to specific areas of scientific research, rather than specific processing operations.

Consent in the GDPR

Article 6(1)(a) GDPR sets free, informed and specific consent as the default condition for the lawful processing of personal data. Furthermore, Articles 6(1)(a) and 9(2)(a) define consent as the default lawful condition for the processing of personal data and sensitive categories of personal data. In this context, Recital 32 states that consent must be active and not obtained via silence or lack of objection.

Another regulatory policy option involves clarifying the conflict between consent under the GDPR and that under the Clinical Trials Regulation (CTR), given that the current European Data Protection Board (EDPB) has identified a conflict between the definitions of consent under these legal instruments. In 2019, the EDPB ruled that a distinction should be drawn between consent for primary purposes (that is, purposes associated with the clinical trial) and content for secondary purposes (that is, research-related purposes). A further distinction can be made under primary purposes between consent for 'safety and reliability' purposes and consent for scientific purposes ('pure research activity purposes'). For safety and reliability purposes, the EDPB held that the processing of personal data could be justified under the CTR pursuant to Article 6(1)(c). Likewise, the processing of sensitive categories of personal data for safety and reliability can be justified under Article 9(2)(i). For research purposes, there are two possible grounds for consent. First, Article 6(1)(a) could be used in conjunction with Article 9(2)(a) for consent as lawful grounds for processing. Secondly, Articles 6(1)(e) and 6(1)(f) could be used together with Articles 9(2)(i) or (j). Nevertheless, for processing both within and outside a clinical trial protocol, the EDPB noted that the standard imposed for consent by the GDPR is different to that imposed by the CTR. Researchers conducting a clinical trial and wishing to reuse participant data for further research would need to obtain GDPR-compliant consent from the research participants.

The GDPR includes a set of exemptions from the normal obligations that are imposed upon data processors and controllers. These exemptions are subject to positive obligations to implement the appropriate safeguards set out in Article 89(1). Article 89(1) (along with the associated Recitals 157 to 162) provides a broad definition of research-related, public interest and statistical processing. More specifically, scientific and statistical processing is not limited to publicly funded academic research but includes privately funded commercial research. One policy option is to use legislation to restrict the scope of the GDPR as regards the forms of scientific or statistical processing that are permitted, to include public interest processing only. Alternatively, research institutes and scientific societies could work together to develop an appropriate ethical standard to prevent the unethical processing of data. There is also a potential conflict as to whether transferred data becomes pseudonymised data. The former Article 29 Working Party conducted a technical review and concluded that anonymisation could only be achieved where it was impossible to re-identify individuals from their data. However, both Court of Justice of the EU (CJEU) and UK case law have concluded that if the recipient of pseudonymised data does not have access to a mechanism to re-identify the data, that data should be considered anonymised. In light of this contextual requirement, institutions should develop guidelines for determining the ongoing status of transferred pseudonymised data.

Another policy option involves developing guidelines to assist with researchers' GDPR compliance in two further areas. First, institutional guidelines could assist with the management of personal data generated via inferences from non-personal data. These guidelines could include mandatory reporting to institutional review boards upon identifying personal data, revising ethics approval, and introducing data protection by design where personal data may be inferred. Second, while the GDPR facilitates the transnational transfer of data within the EU, the normative conditions for transnational transfers of personal data outside the EU might be subject to regulatory ambiguity. This ambiguity might negatively affect the establishment and conducting of transnational collaborative scientific projects involving both EU and non-EU research institutions. As transnational scientific collaboration is key to scientific endeavour and a key driver of innovation, the development of institutional guidelines on transnational data transfer is highly recommended.

There are also specific data subject rights that are ambiguous. First, there are ambiguities inherent within the data subject rights under the GDPR, in particular the right to opt out of automated profiling. Some scientific projects may use big data analytics and machine learning to categorise and profile different data subjects. Furthermore, Article 22 does not provide a specific exception for research subjects vis à vis the obligations contained in Article 89(1). The right to opt out does not apply if the data was processed with the data subject's explicit consent, or where a decision is made solely using automated profiling. Nevertheless, this limited scope does not exculpate researchers from liability. Research institutions and ethics committees should therefore develop appropriate standards for using automated profiling. Second, there is a conflict between data subject rights and the EU *sui generis* database protection regime. This regime provides protection for databases where those databases have an original structure. This can conflict with the rights guaranteed for data subjects under the GDPR, such as the rights to access, information and data portability. This regime can also create inconsistencies with the equivalent right of access guaranteed by the Health Insurance Portability and Accountability Act (HIPAA) in the US. Therefore, scientific societies and research institutes should work collaboratively to determine appropriate limits on the database protection right.

Furthermore, if derogations are introduced under Article 9(4), EU national legislatures should work together to ensure consistency. National data protection authorities (DPAs) and research institutes can play leading roles in facilitating this consistency-seeking harmonisation. Another potential point of ambiguity within the GDPR is with respect to the accuracy principle of data processing under Article 5(1)(d). The translation of Article 5(1)(d) varies with the differing interpretations of the GDPR. In the face of this ambiguity, institutions should develop relevant standards of correctness or accuracy for different fields of research.

2. Procedural options

The procedural options include policy options aimed at enhancing processes and established practices relating to data protection in the context of scientific research. These processes and practices include data management practices, technical standards for privacy-preserving data sanitisation, and technology-assisted regulatory compliance. Installing robust data management practices for researchers can facilitate the successful implementation of the GDPR within the scientific community. Research institutions, academic societies and professional associations should be responsible for the development, promotion and enforcement of such practices. On the basis of the rights and obligations of the GDPR, key focus areas should be data integrity and data security. Furthermore, accountability for the adverse effects of poor data quality should be ensured through transparent data management practices.

Compliance with GDPR data management requirements has emerged as one of the issues most likely to affect scientific research practices. Over the past decade, scientific research has become increasingly reliant on the collection, storage, analysis and re-use of personal data. Such data now includes not only conventional health-related data but also new sources of data, including data generated by research participants themselves through smartphone apps, wearable devices or social media activities. As a result, there is increasing awareness of the need to articulate a coherent set of governance principles to handle the specific challenges raised by the data-centric nature of present-day scientific research. Numerous research stakeholders are developing data governance guidelines that would align existing regulatory instruments with ethical safeguards covering new data-related practices in the scientific space. According to this frame, oversight should be inclusive, anticipatory, innovative and proportionate.

The GDPR recognises and promotes the privacy-enhancing effect of techniques such as data pseudonymisation and anonymisation. The development of technical standards, such as a check-list of specific data elements whose removal ensures pseudonymisation or anonymisation under the provisions of the GDPR might be a viable strategy to achieve this aim. Professional associations and academic societies could play a leading role in developing such harmonised technical standards. GDPR compliance might not necessarily be straightforward for research institutions and individual scientists. Therefore, the development of assistive software-based tools would be a proactive measure to facilitate GDPR-compliance and reduce the administrative burden on researchers. Finally, software tools ensure that the obligations of the GDPR are met in a way that is consistent with the requirements of consent in associated instruments (such as the CTR). This finding may hold particularly true if administrative management tools are used in coordination with digital consent management service tools.

3. Transitional capacity-building options

The transitional capacity-building options include the promotion of activities aimed at disseminating knowledge and building the capacity needed by research institutions to rise to the challenges of GDPR compliance. These activities are defined as transitional because their implementation is instrumental to facilitating the transition to the new GDPR regime.

As the GDPR affects the work of the research community, it is essential that researchers and other professional staff working with personal data understand the basic principles of the GDPR and its implications for data protection in the research context. Some European universities now offer short training courses, also including online courses to train researchers to become familiar with GDPR. Training activities should be targeted and time efficient. Moreover, it will be essential to make relevant information easily accessible and comprehensible for researchers from different fields and disciplines. In this context, the use of knowledge visualisation techniques constitutes a promising approach. Knowledge visualisation is a methodology that can be used to translate complex, linear policy documents into more accessible, interactive formats that can reduce cognitive load and foster learning. GDPR training should also be mandatory for all scientific staff working with personal data to ensure participation in these activities. Such training could, for example, form part of designated training for new scientific staff upon arrival, also including early-career researchers. As the study findings attest, there is uncertainty about the administrative resources required for GDPR compliance. While there is a plausible expectation that smaller research institutions might be disproportionately affected by the requirements for GDPR compliance compared with larger institutions, this hypothesis requires empirical verification. Post hoc empirical studies are critical for comprehensively assessing the impact of the GDPR on scientific research. For this reason, European funding agencies might consider opening targeted calls for proposals, possibly within the framework of existing EU research and innovation programmes.

The study findings revealed ambivalence and mixed perceptions in the scientific community about the potential obstacles and burden imposed by the GDPR on research. This suggests that there is a need to monitor attitudes in the scientific community towards the GDPR continuously in a systematic and longitudinal manner. This continuous monitoring could be implemented using social science methods such as surveys, and qualitative and ethnographic research methodologies. In particular, this research might be useful to identify and possibly address persisting misconceptions about the GDPR. The output of this research could be targeted at information campaigns and educational programmes such as those proposed above. There is a clear need to engage in communication with the public with respect to the collection and use of personal data in the context of scientific research. One strategy might involve developing public awareness campaigns to strengthen public understanding of the individual rights and obligations that are reinforced by the GDPR and raise awareness with respect to the distinctions between scientific and marketing research.

This document is based on the STOA study on 'How the General Data Protection Regulation changes the rules for scientific research' (PE 634.447) published in July 2019. The study was led by Effy Vayena, Health Ethics and Policy Lab, ETH Zurich, at the request of the Panel for the Future or Science and Technology (STOA) and managed by the Scientific Foresight Unit within the Directorate-general for Parliamentary Research Services (DG EPRS) of the European Parliament. STOA administrator responsible: Mihalis Kritikos.

stoa@ep.europa.eu (contact)

http://www.europarl.europa.eu/stoa/ (STOA website)

www.europarl.europa.eu/thinktank (internet)

http://epthinktank.eu (blog)