Participatory Approaches to a New Ethical and Legal Framework for ICT

PANELFIT

D2.1 Issues and gaps analysis on informed consent in the context in ICT research and Innovation

Dissemination level

PU	Unrestricted PUBLIC Access – EU project	X
PP	Project Private, restricted to other programme participants (including the Commission Services) – EU project	
RE	Restricted to a group specified by the consortium (including the Commission Services) – EU project	
CO	Confidential, only for members of the consortium (including the Commission Services) – EU project	

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

Document Information

Grant Agreement n°	788039	
Project Title	Participatory Approaches to a New Ethical and Legal Framework for ICT	
Project Acronym	PANELFIT	
Project Coordinator	UPV/EHU	
Document Responsible Partner	VUB	paul.de.hert@uvt.nl
Document Number	D2.1	
Document Title	Issues and gaps analysis on informed consent in the context in ICT research and Innovation	
Dissemination Level	PU	
Contractual Date of Delivery	30/09/2019	

Partners involved in the Document

N°	Participant organisation name (short name)	Acronym	Check if involved
1	Universidad del País Vasco/Euskal Herriko Unibertsitatea	UPV/EHU	x
2	Fonden Teknologiradet	DBT	
3	Vrije Universiteit Brussel	VUB	x
4	Oesterreichische Akademie der Wissenschaften	OEAW	x
5	Goethe Universität. Frankfurt am Main	GUF	x
7	European Citizen Science Association (ECSA)	ECSA	x
8	European Network of Research Ethics Committees	EUREC	x
9	Consejo Superior de Investigaciones Científicas	CSIC	x
10	Centro per la Cooperazione Internazionale/Osservatorio Balcani Caucaso Transeuropa	CCI/BCT	
12	EVERIS SPAIN, S.L.U.	EVERIS	x
13	Unabhängiges Landeszentrum für Datenschutz AöR	ULD	x

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

Contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

General Introduction

This document corresponds to Deliverable 2.1, as described in the Amended version of Grant Agreement number: 788039 — Participatory Approaches to a New Ethical and Legal Framework for ICT (PANELFIT) (AMENDMENT Reference No AMD-788039-12). According to page 6 of Annex 1, part A (List of Deliverables)

Deliverable 2.1 is entitled "Issues and gap analysis on data commercialization in the context of ICT research and innovation". It corresponds to "**Task 2.1. To prepare an issues and gaps analysis on the regulation of informed consent and decision-making tools in the context of ICT research and innovation", which reads:**

> "WP2 will examine the national and international regulation of informed consent and produce a report that will be incorporated to the Critical Analysis document. To this purpose, they will benefit from the workshops and other engagement activities promoted by the project. Suggestions on improvements of the current legislator framework will be incorporated into this document."

The draft of the Report was elaborated according to the methodology described in the Description of the Action (DoA) while addressing "**Task 2.3. To organize a workshop with experts on the issues at stake and to publish its main results"**, which reads:

> "A workshop devoted to informed consent will be organized by VUB. It will last two working days. 12 worldwide experts will be invited to participate. They will all have to produce papers that will be circulated among participants at least two weeks before the event (…). VUB will promote the edition of these materials in JCR indexed journals. They will be available via open-access."

Accordingly, WP2 organized a workshop on the ethical and legal issues regarding ICT data protection in Bilbao, Spain, in June 2019. It was aimed at uncovering issues that have not been adequately addressed by the new EU regulation and proposing alternative solutions.

During this workshop, the participating members from the PANELFIT consortium and several distinguished external experts (from Belgium, United Kingdom, Germany, Poland, Luxembourg, Italy, Austria, all doctors in law or in related areas with ample publications and considerable research

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

experience in the field)[1] identified the main legal issues and gaps related to informed consent in the currently existing legal EU framework on data protection. They also proposed some feasible **measures to fill in the gaps and solve open issues.** Furthermore, some of the attendants produced extremely interesting academic papers on the role of consent and legitimate interests as legal bases in research, Ethical Impact Assessment in the research sector, transparency rights in research, the notion of scientific and statistical research in the GDPR, and the protection of vulnerable research subjects in the GDPR. These contributions have now been already published in a special number of the leading international legal journal *Computer Law and Security Review* (in open access via this link: https://www.sciencedirect.com/journal/computer-law-and-security-review/vol/37/suppl/C) promoted by PANELFIT WP2. All this information greatly stimulated the work of the WP2 members, who produced a first draft of the Issues and Gaps Analysis in February 2020.

This first draft was shown to twenty additional experts who attended a common workshop held in Madrid between 2 and 4 March 2020. All of them were given the chance to provide feedback on the document and valuable comments were gathered. Based on this feedback, a renewed version of the document was built during March and April 2020. This second version was then reviewed by an expert who participated in the first round of the Extensive Public Consultation in mid May. His comments were used to improve that second version and build the final version of this Deliverable, which was finally sent to the European Commission.

[1] Professor Charles Raab, University of Edinburgh; Professor Dr. Paul Quinn, Vrije Universiteit Brussel; Prof. Dr. Ben Wagner, Wien University; Dr. Denise Amram, Sant'Anna School of Advanced Studies of Pisa; Dr. Rossana Ducato, Université Catholique Louvain La Neuve / Université Saint Louis; Dr. Arianna Rossi, University of Luxembourg; Dr. Jedrzej Niklas, London School of Economics; Dr. Dara Hallinan, FIZ Karslruhe.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

1. Executive summary

The objective of this analysis is not to list all existing issues and gaps related to informed consent and data protection in ICT research, but rather to highlight the most pressing and recently emerging issues and gaps.

In particular, this Analysis focuses on several key issues, such as: the problematic role of informed and free consent in data processing for research purposes; definitory problems; consistency issues across the GDPR; and the most sensitive areas of data processing for research:

- As regards the role of *consent*, this Analysis focusses on the legal and ethical issues for data controllers who adopt more than one legal basis; the alternative between consent and other legal bases (legitimate interests or public interests, in particular for research purposes); how the consent can be really free and informed and how transparency duties can be effective also in cases of significant power imbalance between data subjects and data controllers;
- As regards the *definitory* and *consistency issues*, this Analysis focusses on the problematic interplay between the notion of anonymous data and specific cases of pseudonymized data in particular in the field of ICT research; but also on the problematic definition of "scientific" and "statistical" research in the GDPR.
- As regards the most *sensitive areas*, this Analysis focusses in particular on the processing of special categories of personal data for research purposes and the lack of Member States legislation implementing the legal basis for processing special categories of data under Article 9(2)(j), the protection of vulnerable research subjects, the protection of data of deceased subjects; the role of "other" persons that are not data subjects.

As a summary to the complete analysis carried out below, the gaps and issues identified are listed below with a reference to the corresponding section between parentheses.

- Issue: the important but often problematic role of free and informed consent for data processing in research. (1.1)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

- Issue: the risk of further purposes for processing personal data in case of research. (1.1.4)
- Issue: the often-blurred distinction between the notion of anonymous and the notion of pseudonymous Data (1.2)
- Issue: national Differences on regulating data processing for research purposes (1.3)
- Issue: the issue of legal bases for data processing for research purposes, i.e. public vs legitimate interest and the discrepancies in their choice based on the public or private nature of research entities (1.4.1)
- Issue: the protection of personal data of deceased data subjects (1.6)
- Issue: The role of the Data Protection Impact Assessment for research purposes (1.7)
- Gap: The role of other individuals (different from the main data subject) that could also be affected by the data processing (1.1.5)
- Gap: the lack of Member States law to process data under Article 9(2)(j) (1.4.2)
- Gap: A specific protection for Vulnerable Data Subjects in research (1.5)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

1.1 Consent: first or last choice?

Context and legal background

It is unclear whether consent should always be the preferred legal basis for data processing activities, regardless of the area or field where the processing shall take place. Article 6(1) of the GDPR prescribes that data processing must have at least one legal basis, but it does not oblige data controllers to prefer one over another. For example, the Irish supervisory authority, the Irish Data Protection Commission (IDPC), has indicated that "there is no hierarchy or preferred option within this list, instead each instance of processing should be based on the legal basis which is most appropriate in the specific circumstances"[2]. In particular, regarding the importance of consent, the IDPC indicates that "it is important to note that 'consent', whilst perhaps the most well-known, is not the only legal basis for processing – or even the most appropriate in many cases"[3].

Regarding the limited case of clinical trials, the EDPB, in its Opinion n. 3/2019, clearly states that "consent will not be the appropriate legal basis in most cases, and other legal bases than consent must be relied upon (see below alternative legal bases). Consequently, the EDPB considers that data controllers should conduct a particularly thorough assessment of the circumstances of the clinical trial before relying on individuals' consent as a legal basis for the processing of personal data for the purposes of the research activities of that trial."[4].

However, the EDPS has clarified that consent can be an adequate legal basis for research in general but that "there may be circumstances in which consent is not the most suitable legal basis for data processing, and other lawful grounds under both Articles 6 and 9 GDPR should be considered"[5]. However, the EDPS clarifies that even where consent is not appropriate as a legal basis under GDPR, *informed consent* by a human research participant could still serve as an

[2] Irish Data Protection Commission, 'Legal Bases for Processing Personal Data' (Data Protection Commission 2019) Guidance Note 2 <https://www.dataprotection.ie/sites/default/files/uploads/2019-12/Guidance%20on%20Legal%20Bases_Dec19_1.pdf> accessed 30 March 2020.
[3] ibid.
[4] 'Opinion 3/2019 Concerning the Questions and Answers on the Interplay between the Clinical Trials Regulation (CTR) and the General Data Protection Regulation (GDPR) (Art. 70.1.b))' (European Data Protection Board 2019) paras 20–21 <https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_opinionctrq_a_final_en.pdf> accessed 30 March 2020.
[5] 'A Preliminary Opinion on Data Protection and Scientific Research' (n 17) 20

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

'appropriate safeguard' of the rights of the data subject. Under which conditions such informed consent might be deemed an appropriate safeguard is still unclear.[6]

1.1.1 Is there more than one legal basis for data processing?

Context and legal background

Article 6(1) states: "Processing shall be lawful only if and to the extent that *at least one* of the following applies" (emphasis added)[7]. The inclusion of 'at least one' implies that more than one legal basis is acceptable. In this regard, the Spanish supervisory authority, the Agencia Española de Protección de Datos Personales, has stated that if a data processing activity was based on consent and the data subject revokes it, the data controller can continue with the data processing activity if there is another legal basis to do so[8]; the IDPC arrived at the same conclusion[9]. The EDPB confirms these interpretations made by national supervisory authorities and states that controllers cannot swap between legal bases after the data processing activities have started[10]. Nevertheless, all of these interpretations are not directly related to research and the issue, therefore, remains open.

Issue

As noted above, the ambiguity in the GDPR means that it is not clear across Europe whether it is legally acceptable to have more than one legal basis for data processing (e.g. consent and legitimate/public interest) in the case of research. In other words, should researchers always choose (only) one legal basis, or could they opt for more than one legal basis? In countries where national supervisory authorities have taken a stance on this matter, researchers have clear guidelines. However, the question remains open at a European level, in particular for the cases of cross-border research.

The Article 29 Working Party (WP29) states that "[i]t is important to note here that if a controller chooses to rely on consent for any part of the processing, they must be prepared to respect that choice and stop that part of the processing if an individual withdraws consent.

[6] See, EDPS, Preliminary Opinion, 18-19.
[7] Emphasis added.
[8] 'Protección de Datos: Guía Para El Ciudadano' (Agencia Española de Protección de Datos Personales) 21 <https://www.aepd.es/sites/default/files/2019-10/guia-ciudadano.pdf> accessed 30 March 2020.
[9] 'Legal Bases for Processing Personal Data' (n 1) 10.
[10] EDPB, Guidelines 05/2020 on consent under Regulation 2016/679, Adopted on 4 May, 2020, 23-24.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

Sending out the message that data will be processed on the basis of consent, while actually some other lawful basis is relied on, would be fundamentally unfair to individuals."[11] This has been later upheld by the EDPB[12].

EDPB Opinion 3/2019, rephrasing Article 17(1)(b), affirms that in cases of consent withdrawal, data should be erased "if there is no other lawful basis justifying the retention for further processing", and that this latter situation would be limited to clinical trials.

Therefore, the general rule seems to be that a research entity might not be able to switch legal bases in the event that one basis is no longer available, but the exception would be in certain situations where an authoritative body, such as a national supervisory authority or the EDPB, has issued an opinion, guidance or ruling on the matter that authorizes the continuation of data-processing activities relying on another previously and timely informed legal basis. In effect, research entities would not be able to continue processing data if the activities were based on consent and the data subject revokes it – unless they can rely on another legal basis and have supporting arguments, such as authoritative interpretation, to do so.

Relevance and impact on R&I

Research entities must determine which legal basis, or bases, shall be used in connection with the intended research, conduct DPIA and Prior consultation with Supervisory Authority (Art. 35, 36 GDPR) if necessary, and prepare to respond should their chosen legal basis be removed (e.g. consent is withdrawn). Moreover, research entities, before processing itself starts, will have to review the existing regulatory framework applicable to their project to determine what legal basis can be used for processing of personal data for research purposes, as well as identifying existing supporting opinion from an authoritative body.

Mitigation measures

Clear and precise guidelines are needed to resolve this uncertainty around the possibility of continuing a research project upon the disappearance of the legal basis used. This is required particularly because some national supervisory authorities have already taken a stance on the matter, meaning there are disparities between Member States. Specific EDPB guidelines could

[11] 'Guidelines on Consent under Regulation 2016/679' (Article 29 Working Party 2018) WP259 rev.01 23.
[12] 'Guidelines 05/2020 on consent under Regulation 2016/679' (n 10), 24.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

provide the necessary clarification. Alternatively, codes of conduct for research entities might help mitigate the issue by taking a conservative approach to the matter, and instructing that data-processing activities should cease upon the disappearance of the selected legal basis, and offering guidelines for how research entities should respond to data subjects exercising their rights in such cases.

1.1.2 Easy obtainment of consent versus freedom of consent

Context and legal background

Consent as a legal basis for the processing of personal data, particularly in research, is very widespread and fundamental to legitimising a large proportion of important research. Consent is defined in Article 4(11) of the GDPR as "any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she (…) signifies agreement to the processing of personal data relating to him or her".

Article 7 dictates the requirements for consent, and WP29 has provided guidance regarding how to comply with these. As such, consent must be granular, requested in a clear and comprehensible manner, easily withdrawable and free (e.g. the provision of a service should not be conditional on the consent for processing data that are not necessary for the provision of that service)[13].

Issue

Article 7(4) and Recital 43 acknowledge that there might be an imbalance between the data subject and the data controller, particularly when the controller is a public authority. Often, research entities are public authorities (e.g. public hospitals conducting medical research, universities conducting surveys) or large entities (e.g. private universities). The collection of consent might also appear to be related to the provision of a service, for example when universities give discounts or awards for participation in surveys, or when hospitals conduct research within the provision of healthcare services. Should these cases be considered situations of imbalance in which consent is not free, and therefore other legal bases (e.g. legitimate interest, public interest) should be relied upon? In order to overcome the potential imbalance

[13] 'Guidelines on Consent under Regulation 2016/679' (n 9); 'Guidelines on Transparency under Regulation 2016/679' (Article 29 Working Party 2018) WP260 rev.01.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

between data subject and controller, the latter must provide all relevant information to help in the decision-making process of the data subject.

In the age of e-commerce and the Internet of Things (IoT), obtaining consent has become extremely easy (e.g. ticking a checkbox on a webpage), and more and more experts are questioning the value of consent as a means to protect individuals.[14] Also, verifying that all requirements for consent stated in Article 7 are respected when obtaining it is difficult, or sometimes even impossible. Consequently, people may grant consent without being fully aware of the implications of doing so, rendering their consent meaningless.

Relevance and impact on R&I

Consent must be given as a clear, affirmative act that establishes a freely given, specific, informed and unambiguous indication of the subject's agreement to the processing of their personal data. Researchers may not always be able to fulfil these requirements, however, due to their lack of knowledge regarding how consent should be collected, as well as the lack of tools to convey, in a compliant manner, all the information prescribed by Article 13 of the GDPR. It is necessary, for example, to provide accurate and precise information on consent in a language that can be widely understood. In practice, this means that researchers should make it more explicit to the individual that they have to agree to their participation in the research, as well as to any data-processing activities resulting from the project[15].

Mitigation measures

While there is guidance on how to address this issue in general, provided by WP29, specific guidance from the EDPB or EDPS for researchers and research institutions on how to correctly

[14] Damian Clifford, Inge Graef and Peggy Valcke, 'Pre-Formulated Declarations of Data Subject Consent – Citizen-Consumer Empowerment and the Alignment of Data, Consumer and Competition Law Protections' (Social Science Research Network 2018) SSRN Scholarly Paper ID 3126706 <https://papers.ssrn.com/abstract=3126706> accessed 19 February 2019; Gabriella Fortuna-Zanfir, 'Forgetting about Consent. Why the Focus Should Be on "Suitable Safeguards" in Data Protection Law' in Serge Gutwirth, Ronald Leenes, Paul De Hert (ed), *Reloading Data Protection* (Springer 2014); S. van der Hof, 'I Agree... Or Do I? A Rights-Based Analysis of the Law on Children's Consent in the Digital World' (2017) 34 Wisconsin International Law Journal 409; Bart W Schermer, Bart Custers and Simone van der Hof, 'The Crisis of Consent: How Stronger Legal Protection May Lead to Weaker Consent in Data Protection' (2014) 16 Ethics and Information Technology 171.

[15] Nijhawan, L. P., Janodia, M. D., Muddukrishna, B. S., Bhat, K. M., Bairy, K. L., Udupa, N., & Musmade, P. B. (2013). Informed consent: Issues and challenges. Journal of Advanced Pharmaceutical Technology & Research, 4(3), 134.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

convey the information would be useful. This should provide not just information, which is already available, but also tools to help with the collection of consent in a GDPR-compliant manner.

1.1.3 Consent for participation in research projects, and consent as a legal basis

Context and legal background

WP29 states that "[w]hen consent is the legal basis for conducting research in accordance with the GDPR, this consent for the use of personal data should be distinguished from other consent requirements that serve as an ethical standard or procedural obligation"[16]. This criterion is further clarified by the EDPB in Opinion 3/2019, which states that consent for participation in research projects and consent as a lawful basis should be separate and different[17]. Recently, the EDPS has also remarked that "[t]here is clear overlap between *informed consent* of human participants in research projects involving humans and *consent* under data protection law. But to view them as a single and indivisible requirement would be simplistic and misleading."[18] Therefore, it seems that the consensus is to always separate the collection of both types of consent.

Issue

Informed consent is an appropriate legal basis under the GDPR for processing data in many situations, including research, but this must be distinguished from the consent research subjects are asked to provide to be involved in a research project. Data subjects must understand that they are granting consent for two separate matters, and that their consent can be revoked independently for the research, i.e. the activities carried out by a scientist in an attempt to make scientific progress, and the data-processing activities, i.e. the auxiliary activities that help a scientist analyse and digest the information collected from the actual research. Otherwise, research entities could be operating under the impression that both forms of consent have been duly and lawfully collected when this has, in fact, not happened.

Relevance and impact on R&I

'Guidelines on Consent under Regulation 2016/679' (n 9) 28.

[17] EDPB, Opinion 3/2019 concerning the Questions and Answers on the interplay between the CTR and the GDPR (Article 70.1.b), Adopted on 23 January, 2019, 5-6.

[18] EDPS, Preliminary Opinion, 20.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

The legal relevance and implications of these two forms of consent are very different. For example, data subjects might withdraw their consent only for the processing of personal data, but not for participation in the research project. Therefore, the research entity has to be able to comply with the data subject's request and determine with which part of the study each data subject agrees.

Mitigation measures

The differences between consent for participation in research projects and consent as a legal basis for data processing must be clearly emphasized by the institution carrying out the research, and at the point at which they request and collect consent. In the event that the entity intends to use a single consent form, then a clear record of what each participant has consented to in connection with the project should be maintained. In this regard, the development of codes of conduct, with best practices and template forms, should be encouraged to provide research entities, especially smaller ones and those less experienced in dealing with these issues, with the tools necessary to achieve a higher level of compliance with the accountability principle.

1.1.4 Foreseeing all use cases versus broad consent

Context and legal background

Article 5(1)(b) requires that legal bases for data processing should be always explicit, legitimate and determinate. However, Recital 33 indicates that research entities can obtain broad consent for further data-processing activities under the following terms: "It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose." Consequently, this apparent broad consent can only be given for specific areas of future research that are related to the original research for which consent was given. Also, this is only possible when in keeping with recognized ethical standards.

Issue

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

WP29 states that Recital 33 does not constitute a waiver for the data controller regarding the requirements that the collected consent must meet. In particular, WP29 indicates that "it should be noted that Recital 33 does not disapply the obligations with regard to the requirement of specific consent. This means that, in principle, scientific research projects can only include personal data on the basis of consent if they have a well-described purpose. For the cases where purposes for data processing within a scientific research project cannot be specified at the outset, Recital 33 allows as an exception that the purpose may be described at a more general level."[19]

Therefore, research entities must foresee all possible purposes for data processing for research purposes. Here, the problem lies in how the research entity should inform the data subject about the future purposes. For example, should research entities establish databases with data subjects' contact info (phone, address or email) and in case of further processing, inform the data subject?

Another problem is the lack of clarity about the recognized ethical standards to be applied: should these ethical standards be established at the national level or, for example, by an association of researchers? Furthermore, how should they be established: through certification measures, for example? And who should check these standards are being respected? Can data protection authorities assess the respect of ethical guidelines? If not, who should do this?

Relevance and impact on R&I

In light of these requirements, a research entity might choose to obtain very broad consent. However, this would be ethically questionable and in contradiction with the principles of lawfulness, transparency and fairness at Article 5(1)(a), and with the purpose limitation principle in Article 5(1)(b), thus rendering the legal basis for data processing unlawful. This situation is even more problematic if special categories of personal data are involved, as WP29 considers that "the flexible approach of Recital 33 will be subject to a stricter interpretation and requires a high degree of scrutiny"[20].

Mitigation measures

[16] 'Guidelines on Consent under Regulation 2016/679' (n 9) 28–29.
[17] ibid 29.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

Research entities that intend to conduct further research, while relying on the original consent given, should inform the data subjects of this possibility, following the indications of Recital 33, and keep track of the purposes for which consent was provided. As proposed by WP29, "having a comprehensive research plan available for data subjects to take note of, before they consent, could help to compensate a lack of purpose specification. This research plan should specify the research questions and working methods envisaged as clearly as possible."[21]

In this sense, transparency constitutes the best mechanism to address the lack of certainty at any stage, as suggested by WP29: "a lack of purpose specification may be offset by information on the development of the purpose being provided regularly by controllers as the research project progresses so that, over time, the consent will be as specific as possible. When doing so, the data subject has at least a basic understanding of the state of play, allowing him/her to assess whether or not to use, for example, the right to withdraw consent pursuant to Article 7(3)."[22] For every processing of data that is not in line with the originally declared purposes, consent needs to be obtained again, unless another legal basis applies - as outlined in Section 1.1.1 - and this has to be made clear. In this regard, research entities should deploy 'talkback' systems that allow data subjects to be contacted, to determine if they permit the new purpose and to obtain new consent before attempting to use other legal bases.

Guidelines for researchers should clarify that, when obtaining consent, the purposes should not be too general or broad - but at the same time, in cases of further purposes, safeguards should be adopted (e.g. clear information notice updates, easy opt-out mechanisms, etc.). Further clarification on this matter by the EDPB, following WP29 guidelines, should be provided specifically for research institutions and to aid the interpretation of Recital 33. A clear code of conducts for researchers, which clarifies the ethical standards to follow according to Recital 33, would also help.

1.1.5 The issue of other data subjects

Context and legal background

[18] ibid.
[19] ibid.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

The definition of a data subject at Article 4(1) of the GDPR refers to the person to whom the data are related. They are the only person who needs to be informed about data processing according to Articles 13 and 14, and the only person to take into account when selecting the legal basis for processing personal data under Article 6 (e.g. in case the data controller decides to process data on the basis of consent, this is the only person from whom consent needs to be requested).

However, the definition of personal data in Article 4(1) is broad and, therefore, it is possible that when collecting personal data about a particular individual, the data controller is also collecting personal data from other natural persons related to that individual[23] (e.g. relatives of the data subjects).

Gap

Often, the personal data of one data subject has informational value regarding other natural persons;[24] for example, the genetic data of a grandparent will also tell researchers something about their grandchildren.[25] Predictive analytics allow the inference of an enormous amount of data not only about the data subject, but also about other related subjects. This means that the processing of personal data will not only affect the data subject, but also other data subjects.

Relevance and impact on R&I

In this situation, research entities would incur the obligation prescribed by Article 14 and would have to inform all other related data subjects of the collecting of their personal data, and how is it processed. However, in many cases, researchers will be unaware of these duties, or incapable of contacting these individuals. In particular, data controllers would have to determine which legal basis the processing of data from the other data subjects is used. It is also possible that researchers would have to delete that data if no legal basis can be relied upon.

Mitigation measures

[23] Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) 2 Columbia Business Law Review <https://papers.ssrn.com/abstract=3248829> accessed 18 December 2018.
[24] Gianclaudio Malgieri and Giovanni Comandé, 'Sensitive-by-Distance: Quasi-Health Data in the Algorithmic Era' (2017) 26 Information & Communications Technology Law 229.
[25] See, in particular, Mark Taylor, *Genetic Data and the Law: A Critical Perspective on Privacy Protection* (Cambridge University Press 2012) 107-112.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

Unless the research entity can demonstrate the application of one of the exemptions prescribed by Article 14, special attention should be paid to the forms and other documents used to collect data from the data subjects actively participating in the research project, and to avoid collecting information that could result in the identification of other related data subjects. Specific guidelines should be created to clarify the conditions for processing the data of these data subjects, and include safeguards to be adopted to guarantee the right to privacy and data protection of all data subjects.

1.1.6. Ineffective information duties

Context and legal background

Articles 13 and 14 of the GDPR state which information should be communicated to data subjects regarding the data processing activities. Further, Article 12(1) clarifies that such communication should be "concise, transparent, intelligible and easily accessible form, using clear and plain language".

Issue

Information notices are often ineffective at actually informing data subjects of the risks and details of data processing. In practical terms, there are few guidelines on how this should be done, and how supervisory authorities can assess this (e.g. intelligibility, easy accessibility, clearness and plainness of language). Article 12(7) encourages providing information via icons, and the European Commission has the authority to suggest such icons. However, to date no steps have been taken in this direction.

Relevance and impact on R&I

Providing information on processing that is easy to understand and overseeing of the potential risks for average data subjects is extremely urgent and relevant for data processing in research: providing data subjects with precise and clear information is not only a general ethical requirement; it can also be considered a necessary safeguard to adopt in case of research data processing according to Article 89 (Section 1.5).[26]

[26] Arianna Rossi and Gabriele Lenzini, 'Transparency by Design in Data-Informed Research: A Collection of Information Design Patterns' (2020) 37 Computer Law & Security Review 105402 See also, Rossana Ducato (18).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

Mitigation measures

Clearer guidelines should be released to make information notices more effective for data subjects in research. In addition, the EC could develop the list of icons recommended in Article 12(7). Moreover, issuing templates for data-flow diagrams, which show what will happen with the data, alongside multi-layered documents with the information required by Articles 13 and 14 as applicable, may help data subjects understand how the research entity will handle the data.

1.2 Anonymous versus pseudonymous data

Context and legal background

The notion of **personal data** is fundamental to discussions around informed consent. This is the material scope of the GDPR, according to Article 2(1), and defined in Article 4(1) as "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".

While **anonymous data** is not expressly defined in the GPDR, Recital 26 states that it can be considered as "information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable".

Pseudonymous data is, according to Article 4(5), information that "can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person".

Recital 26 states that "[p]ersonal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information, should be considered to be information on an identifiable natural person". This is because the data controller, or any other third party, could connect the pseudonymized information with the additional information

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

in order to connect the former with an identified or identifiable individual. However, as Recital 26 indicates, "[t]o ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments".

Issue

Both anonymous and pseudonymous data could start as personal data before being subject a processing activity that made them anonymous or pseudonymous data, respectively; nevertheless, it is also perfectly possible for anonymous data to start as anonymous if certain precautions are taken when collecting it.. In recent discussions, several authors[27] suggest that, particularly in the field of research, pseudonymous data can often also be considered anonymous data, and vice versa. In their view, it is unclear when pseudonymized data can actually be considered anonymized data, since the state of the art in data aggregation is constantly evolving and, therefore, anonymized data can be more easily transformed back into personal data.[28] This is leading to a potential risky situation where data is processed as if it were anonymous – and thus without the necessary safeguards – when it is actually pseudonymous and can be linked, with increasing ease, to additional data.[29]

The following example sets out how this can create problems. Researcher A collects personal data for research purposes and uses pseudonymization techniques, keeping the identifying additional information separate and secure. Then, Researcher B asks Researcher A to share the 'pseudonymized data' for conducting further research. Researcher B receives only de-identified data that lacks any identifiers (i.e. the additional information). But, can we be certain that, for Researcher B, such data are indeed anonymous, as she lacks access to the additional information and thus the means to re-identify the pseudonymous data? In other words, can we be certain

[27] Miranda Mourby and others, 'Are "Pseudonymised" Data Always Personal Data? Implications of the GDPR for Administrative Data Research in the UK' (2018) 34 Computer Law & Security Review 222.
[28] See, in particular, Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques, 4.
[29] See, e.g., Sophie Stalla-Bourdillon and Alison Knight, 'Anonymous Data v. Personal Data - a False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data' (2017) 34 Wisconsin International Law Journal 284, 306; See also Luca Bolognini and Camilla Bistolfi, 'Pseudonymization and Impacts of Big (Personal/Anonymous) Data Processing in the Transition from the Directive 95/46/EC to the New EU General Data Protection Regulation' (2017) 33 Computer Law & Security Review 171.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

that Researcher B is not also a data controller since there is no personal data involved, and that the GDPR therefore does not apply to her processing?

The Information Commissioner's Office (ICO) affirmed[30] that "personal data that has been pseudonymised—e.g. key-coded—can fall within the scope of the GDPR depending on how difficult it is to attribute the pseudonym to a particular individual". In other words, the ICO does not exclude the possibility that pseudonymized data can be considered anonymous data when certain characteristics have been met and, consequently, fall outside of the GDPR and its application. Similarly, Mourby et al.[31] state that there may be circumstances in which data that have undergone pseudonymization within one organisation could be anonymous for a third party.

However, this interpretation creates uncertainty and ambiguity: the same set of data would be personal data (though pseudonymized) for some researchers (e.g. Researcher A in the example above) and anonymous data for other researchers (e.g. Researcher B in the example above). In particular, there might be scenarios in which Researcher B eventually obtains the additional information necessary to re-identify individuals from Researcher A. Would she become the data controller and have to comply with the obligations, such as having a legal basis for data processing, only from that moment?

The Court of Justice of the European Union (CJEU) seemed to disagree with the interpretations by the ICO and Mourby and others: in the *Breyer Case,* the Court affirmed that if there are 'legal means' (or, better, if it is not 'prohibited by law', as the Advocate General affirmed[32]) to associate identifiers to the pseudonymized data of a third party, such data should still be considered personal data[33]. Therefore, it is not only a matter of technical possibilities, but also of legal possibilities.

Relevance and impact on R&I

[30] Information Commissioner's Office (ICO). Overview of the General Data Protection Regulation (GDPR). https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/, accessed 26 August, 2019.
[31] Mourby and others (n 24).
[32] Opinion of Advocate General Campos Sánchez-Bordona, delivered on 12 May 2016 (1) Case C-582/14 *Patrick Breyer v Bundesrepublik Deutschland*, §68.
[33] European Court of Justice, Case C-582/14, Request for a preliminary ruling under Article 267 TFEU from the *Bundesgerichtshof* (Federal Court of Justice, Germany), made by decision of 28 October 2014, received at the Court on 17 December 2014, in the proceedings *Patrick Breyer v Bundesrepublik Deutschland*, 19 October 2016, §§44-46.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

It is very common that one research entity pseudonymizes data, and several other research entities (or private researchers) use such de-identified data in separate processing operations. Both anonymous and pseudonymized data play an important role in organisations and the applicability of the GDPR in research depends on this interpretational issue.

If the supposedly anonymized data turns out to be personal data, then the researcher would need to have a legal basis to process it, such as informed consent from the data subjects. This situation is even more relevant in the case of special categories of personal data (Article 9(1); see Section 1.9), where legal bases and the processing activities themselves are far more detailed and limited.

Moreover, even if researchers want to take a conservative approach from the moment they receive a dataset with supposedly anonymized data, it is unclear which standard, either technical or legal as mentioned above, they should apply to determine whether the information they have under their control is actually anonymous or not. This has implications for other data protection obligations, such as the information obligation prescribed in Article 14 of the GDPR, which – inter alia – requires the data controller to disclose the legal basis used. If the researcher, under the currently uncertain legal framework, decides to rely on technical criteria to determine when the data could become personal data, it might delay the provision of the data for processing activities.

Mitigation measures

We propose three potential courses of action to resolve the challenges around anonymous vs pseudonymous data.

(1) An authoritative body, such as the European Data Protection Board (EDPB), should clarify which criteria should be used to determine if information is properly anonymized or not.

(2) Researchers should conduct anonymisation/pseudonymisation audits when receiving third-party datasets, to assess if the information is personal data or not, whether they need to take further action or not (such as obtaining consent or applying technical and organisational measures).

(3) Researchers should treat any data related to individuals as personal data, regardless of whether or not it can be considered as such under legal or technical standards.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

1.3 Different national legislations on research within Europe

Context and legal background

In the area of data protection for research purposes, and in the special categories of personal data (see Section 1.9), the GDPR allows for different national legislations. In particular, Articles 89(2) and (3) allow Member States to regulate more specific derogations and safeguards with regards to data processing for research purposes. Similarly, according to Article 9(4), Member States may maintain or introduce further conditions, including limitations, with regard to the processing of genetic data, biometric data or data concerning health.

Issue

These differences in national legislation create disparities in the application of data protection rules across Europe, particularly in research. Such disparities may encourage unfair practices, such as 'forum shopping', where research initiatives locate their main research partner in Member States with looser requirements for safeguards regarding data processing for research purposes or special categories of personal data.[34]

Relevance and impact on ICT R&I

This issue is particularly relevant for European research projects (especially, but not only, Horizon Europe projects), as they typically operate in several different Member States, with different national rules, derogations and limitations. This may affect the quality of research and lead to inequalities between research institutions and their staff. For example, certain countries might require more complex technical measures to secure data involved in research while others might be more lenient on the matter. Therefore, certain institutions might be obliged to spend resources in dealing with regulation that their peers in other European jurisdictions might not have to.

[34] See, largely, Kärt Pormeister, 'Genetic Research and Applicable Law: The Intra-EU Conflict of Laws as a Regulatory Challenge to Cross-Border Genetic Research' (2018) 5 Journal of Law and the Biosciences 706.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

It may also reduce or impinge on individual rights[35]. In particular, the data subject should be informed about where their personal data will be processed, as prescribed by Articles 13 and 14 of the GDPR, so that researchers are able provide the data subject with appropriate information about data processing and, in case the processing is based on consent, collect an actual freely, informed and specific consent from data subjects.

Mitigation measures

As possible solutions to this issue, we proposed three courses of action.

(1) A fully harmonized European code of conduct on common safeguards for research (in compliance with Article 89), including for processing special categories of personal data. This should be proposed by the EDPB or by specific European Commission Initiatives, involving all appropriate stakeholders (e.g. national research ethics boards, national associations of researchers, etc.).

(2) The creation of a common legal framework for research purposes, possibly adopting the higher standards of protection found across different EU countries.

(3) Before starting any European research project, the partners should clarify which national legislation they will be following by determining the roles that each partner will have on the project (i.e. the law applicable to the controller, who is presumably the Coordinator of the research consortium, while the rest of the partners will act as processors).

1.4 Data processing purposes

Data controllers can use different legal bases to process data for research purposes. For example, they can use consent (Article 6(1) letter a), legitimate interest (Article 6(1) letter f) or public interest (Article 6(1) letter e). For all these legal bases, issues may arise regarding their interpretation or application. We have identified the following issues related to data processing for these varying purposes: (1) a potential disparity between the legal bases available to public

[35] Chassang, G. (2017). The impact of the EU general data protection regulation on scientific research. ecancermedicalscience, 11.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

and private research entities; and (2) processing of special categories of personal data for research purposes.

1.4.1 Processing carried out in the public interest: potential disparity between private and public research entities

Context and legal background

While informed consent could be considered as the most preferable legal basis for data processing, it is possible to rely on other legal bases.[36] As such, the purpose of research could be stated considered as the legitimate interest of the researcher. Accordingly, if data controllers are not willing or able to obtain consent from all data subjects, they could choose to process personal data under the 'legitimate interest' legal basis. However, the second subparagraph of Article 6(1) clarifies that point (f) of the first subparagraph (legitimate interest) shall not apply to processing carried out by public authorities in the performance of their tasks. Therefore, it is possible that research entities that are public authorities cannot rely on legitimate interest as the purpose for data processing.

The GDPR does not provide a definition of a public authority, nor does it refer to national laws to determine its meaning. Thus, the term should be given an autonomous EU-wide meaning.[37] According to the CJEU, the concept of 'public authority' generally encompasses "entities which, organically, are administrative authorities, namely those which form part of the public administration or the executive of the State at whatever level. This first category includes all legal persons governed by public law which have been set up by the State and which it alone can decide to dissolve; [and] entities, be they legal persons governed by public law or by private law, which are entrusted, under the legal regime which is applicable to them, with the performance of services of public interest and which are, for this purpose, vested with special powers beyond those which result from the normal rules applicable in relations between persons governed by private law".[38]

---

[36] See European Data Protection Supervisor, *A Preliminary Opinion on data protection and scientific research*, 6 January 2020., 18.

[37] See, e.g. Case C- 279/ 12, Fish Legal and Shirley, para. 42 and case law cited therein. See, on this point, Luca Tosoni, 'Article 4(22). Supervisory Authority Concerned' in Lee A. Bygrave and Christopher Docksey (eds), *The EU General Data Protection Regulation (GDPR) - A Commentary* (Oxford University Press).

[38] See, e.g. Case C- 279/ 12, Fish Legal and Shirley, para. 51-52. See too e.g. Case C- 188/ 89, *Foster*, para. 20; Case C- 425/ 12, *Portgás*, para. 24.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

Issue

The differences between the eligible legal bases available to justify data processing might create an unreasonable disparity between public and private research entities, such as universities. As an example, even if they have similar research objectives and adopt similar safeguards, they would be required to use different legal bases for their research. Therefore, this issue can be divided into two questions: (1) which research entities should be considered public bodies; and (2) does Article 6(1) refers exclusively to public authorities, or does it include public bodies?

Regarding the first question, and as noted in the 'Context and legal background' paragraph above, performing data processing with the legal basis of public interest, but with 'special powers', might be enough to consider research entities as public entities under the GDPR. Also, according to the European Data Protection Supervisor (EDPS), research purposes that serve a public interest, e.g. improving the provision of healthcare services, can generally be considered within the public interest.[39]

Even if there were consensus about the 'public interest' status of research undertaken by public universities, however, it might become more problematic to establish the status of research by private universities or private research centres: should they be considered private entities (because this is their legal status under national law, or should they be considered public entities under EU law because they perform a public interest, especially in the context of EU-funded projects? If we accept the first option, private universities or private research entities have the freedom to use 'legitimate interest' as a legal basis, while researchers of public bodies who process personal data for research purposes cannot choose the legal basis of 'legitimate interest', but instead would be forced to use 'public interest' as the legal basis.

This is important, because the different bases places certain limits on what researchers can and cannot do with the data. For example, this disparity becomes evident when considering obligations regarding the exercise of the rights of certain data subjects. The right to object to how one's personal data is used can be limited in cases where the 'public interest' legal basis applies, but not in cases of 'legitimate interest'. In particular, Article 21(6), which explicitly refers to research purposes, states that "[w]here personal data are processed for scientific or historical research purposes or statistical purposes pursuant to Article 89(1), the data subject,

[39] See EDPS, A Preliminary Opinion on Data Protection and Scientific Research, 23.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

on grounds relating to his or her particular situation, shall have the right to object to processing of personal data concerning him or her, **unless the processing is necessary for the performance of a task carried out for reasons of public interest**" (emphasis added). In other words, the exercising of the right to object can have different outcomes, depending on whether the data controller is a public or private research entity (and, as a result, the legal basis they have used for data processing).

Also, according to Article 6(2), Member States may maintain or introduce more specific provisions to adapt the application of the rules of the GDPR with regard to processing for compliance with - inter alia - data processing based on public interest, for example by determining more specific requirements for data processing, and other measures to ensure lawful and fair processing. And, according to Article 6(3), national laws regulating data processing for public interest may contain specific provisions to adapt the application of rules of the GDPR, inter alia, "the general conditions governing the lawfulness of processing by the controller; the types of data which are subject to the processing; the data subjects concerned; the entities to, and the purposes for which, the personal data may be disclosed; the purpose limitation; storage periods; and processing operations and processing procedures, including measures to ensure lawful and fair processing".

This can lead to situations in which some data protection rules and safeguards (at the national level) apply only to research led by public research entities (universities of research centres), not private ones. In other words, private research entities could profit from their private status and choose 'legitimate interest' as a legal basis for processing personal data when they want to avoid national legislations under Article 6(2) or (3), and use "'public interest' when the application of national rules is more convenient to them.

Relevance and impact on ICT R&I

The disparity between the status of different research entities is highly relevant to ICT R&I. In particular, it can lead to the situation where a data subject can object to data processing activities by private research entities acting under legitimate interest, but not in the case of processing by public research entities. Also, there might be situations where national laws regulating data processing for research purposes, on the basis of Article 6(2) and (3), only apply to public

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

research entities and not private ones, meaning the latter would be free to process data on the 'legitimate interest' legal basis.

In the field of ICT research, this is especially relevant: the exercise of the right to object might be even more important if the research is conducted through automated means and through the use of innovative technologies (whose risks might be realized by data subjects only at a later stage, when they would perhaps decide to object). Further, Member States may adopt specific data-protection rules for ICT research that may have disparate applications if they only apply to public research entities.

Mitigation measures

To solve these issues, a twofold course of action is recommended.

(1) Regarding the interpretation of Article 6(1), and following the example of the EDPS (which does not mention legitimate interest among the legal bases for scientific research),[40] the EDPB could clarify that all research entities can process data only under consent or public interest bases, and not legitimate interest.

(2) Otherwise, guidelines from an authority such as the EDPB, or even national supervisory authorities since the differences might occur at a national legislation level, could bridge the differences between research conducted by public and private research entities, regardless of the legal basis used.

1.4.2 Processing of special categories of personal data for research purposes

Context and legal background

For the processing of special categories of personal data (Article 9(1); see Section 1.9), the GDPR provides a legal basis that specifically addresses research (Article 9(2) letter j): 'processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject'. However, the GDPR does not provide

[40] See, EDPS, A Preliminary Opinion on data protection and scientific research, passim.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039.

a definition of what constitutes research and statistical purposes and, as of the date hereof, there is no definitive official interpretation on the matter.

This subject brings up both an issue – the definition of research purposes – and a gap – the lack of Member State law for the application of Article 9(2)(j). These are discussed separately in the following sub-sections.

Issue

The definitions of both 'research purposes' and 'statistical purposes' are unclear. From a layman's perspective, research could be understood as discovering something new out of an analysed dataset and making conclusions that will serve for future benefit, while statistics are a manner of displaying certain information. Regarding the notion of research, the GDPR, in Recital 159, states that it has to be interpreted broadly: 'For the purposes of this Regulation, the processing of personal data for scientific research purposes should be interpreted in a broad manner including for example technological development and demonstration, fundamental research, applied research and privately funded research'. This emphasis on a broad interpretation, together with the lack of a more precise definition for this term, means critical questions remain unanswered. For example, is purely private research (e.g. for marketing purposes) included in this broad definition?

Current draft interpretations by authoritative bodies, such as the EDPS, incline towards a restrictive interpretation that only considers research for purposes that are done with the "aim of growing society's collective knowledge and wellbeing, as opposed to serving primarily one or several private interest"[41]. However, the EDPB has not yet endorsed such a definition and, even accepting this definition, it is a very extensive notion open to many different interpretations. For example, on the other hand, the GDPR fails to provide a definition for the term 'statistical purposes', which places it in a less determined area than 'research purposes'. This lack of clarity is exacerbated by the fact that there is not even a draft opinion from an authoritative body on this matter. As a result, it is not clear whether, for example, data

[41] 'A Preliminary Opinion on Data Protection and Scientific Research' (European Data Protection Supervisor 2020) 12 <https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf> accessed 30 March 2020.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

controllers are allowed to use this exception in possibly unintended cases, such as statistical research for marketing purposes.[42]

Relevance and impact on R&I

The definition of both 'research' and 'statistical', and the scope of such definitions, is greatly relevant for data processing for these purposes. The uncertainty of these definitions could lead to uncertainty in the application of Article 9(2) letter j, on the processing of special categories of personal data, and also in the application of the 'exceptions of purpose' limitation principle for research purposes (Article 5(1) letter b).

Mitigation measures

While draft interpretations of these terms have emerged during our analysis, such as those from the EDPS mentioned above, it is necessary that these definitions are formally issued by the relevant authoritative body, such as the EDPB or EDPS, to resolve the uncertainty around this matter. Moreover, an interpretation of 'statistical' research should be provided alongside an interpretation of 'research'.

Gap

One gap in the processing of special categories of personal data is the lack of Member State laws for the application of Article 9(2)(j) of the GDPR, which states that besides consent it is possible to process special categories of personal data if the 'processing is necessary for (…) scientific or (…) research purposes (…)'. However, that legal basis is conditional on the approval of a "Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific" safeguards. As the EDPS argued, many Member States have not yet approved these laws.[43] Accordingly, researchers might not use that exemption and would be required to ask for explicit consent (Article 9(2)(a)), even though it might be inappropriate in case of power imbalance between researchers and subjects.

At the time of writing, in response to the ongoing COVID-19 pandemic, most Member States have approved emergency pieces of legislation that include the processing of special categories

[42] See Rossana Ducato, 'Data protection, scientific research, and the role of information', (2020) 37 Computer Law & Security Review 105402
[43] European Data Protection Supervisor, Preliminary Opinion, 23.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

of data for research purposes.[44] However, as the EDPB has clarified, these legislations should be strictly limited to the duration of the emergency at hand and prove the necessity and proportionality of any derogatory provisions.[45] The issue of necessity and proportionality is also central to the interpretation of DPIA (Article 35(3)(b)).[46]

Relevance and impact on R&I

The uncertainty around the applicability of Article 9(2)(j), on the processing of special categories of personal data for research purposes, is extremely relevant for ICT research. If the lack of national regulations is an impediment to the reliance on this legal basis by researchers, any data processing activity, done under the scope of a research, that involves special categories of personal data could only be done relying on consent. In this regard, researchers would have to adopt several measures to achieve the standard required for consent to be valid, as discussed in Section 1.1.

Mitigation measures

Member States should approve, as soon as possible, national laws to allow researchers to process special categories of Article 9(2)(j). A high level of harmonization is recommended to avoid disparities and 'forum shopping' (see Section 1.2 above). A comparison of the already approved laws is also necessary. In this regard, an action from a European institution could help in making these regulations homogenous as well as pushing Member States to adopt the regulation that researchers might need to conduct with more freedom their projects.

1.5 Vulnerable data subjects

Context and legal background

As discussed in the previous sections, consent for data processing has to be informed and, in return, the information has to be tailored to the data subject, in particular if the data subject is a

[44] See, as the first case in the EU, the Italian Decreto Legge n.14/2020, Article 14.
[45] European Data Protection Board, 'Statement on the Processing of Personal Data in the Context of the COVID-19 Outbreak. Adopted on 19 March 2020' (2020) 2 <https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_statement_2020_processingpersonaldataandcovid-19_en.pdf>.
[46] See, e.g., Dariusz Kloza and others, 'Towards a Method for Data Protection Impact Assessment: Making Sense of GDPR Requirements' 6.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

vulnerable individual[47]. It is therefore crucial, to properly comply with this obligation, to determine whether or not the data subject is a vulnerable individual.

However, there are different understandings of, vulnerable individuals within Europe. Nor does the GDPR provide a list of vulnerable individuals;[48] Recital 75 merely presents the concept of vulnerable data subjects, with only children mentioned as an example. WP29 guidelines on Data Protection Impact Assessments (DPIAs) mention that 'vulnerability' is a key index to consider when assessing the necessity of conducting a DPIA. Some categories are listed, such as employees and the "more vulnerable segment of the population requiring special protection, such as, for example, the mentally ill, asylum seekers, or the elderly, a patient, or in any case where an imbalance in the relationship between the position of the data subject and the controller can be identified"[49]. The EDPB Opinion 3/2019 only mentions "economically or socially disadvantaged people"[50]. Meanwhile, the Clinical Trial Data Regulation, which only refers to a specifically limited and delicate area of research, considers different categories of vulnerable individuals in research: frail, multiple chronicle conditions, mental disorders, older (Recital 15), incapacitated (Article 10(2)), pregnant or breastfeeding (Article 10(3))[51].

Gap

Considering the regulatory landscape described above, including the interpretations provided by supervisory authorities, it is clear that the notion of a 'vulnerable person' is not unanimously agreed upon and therefore not all vulnerable individuals are equally protected. As indicated, there is no definitive and universal interpretation of this, leaving it up to data controllers to determine when there is a vulnerable person and how their situation is to be addressed, in particular regarding how to properly convey information to them.

As noted above, the 'public interest' legal basis does not guarantee the full exercise of a right to object, or the right to data portability. The EDPB recently clarified[52] that - at least in the area

[47] 'Guidelines on Transparency under Regulation 2016/679' (n 10) paras 14–16.
[48] However, see largely Gianclaudio Malgieri and Jedrzej Niklas, 'The Vulnerable Data Subject' (2020) 37 Computer Law & Security Review.
[49] Article 29 WP, WP248, 9.
[50] EDPB, Opinion 3/2019 concerning the Questions and Answers on the interplay between the CTR and the GDPR (Article 70.1.b), Adopted on 23 January 2019, 6.
[51] Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April, 2014, on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC (text with EEA relevance).
[52] Ibidem.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

of clinical trials, but we can assume by analogy also in similar cases[53] -several vulnerable individuals might not provide free consent to data processing for research purposes. As such, research entities have to rely on other legal bases for data processing (e.g. legitimate interests) that do not allow the exercising of the full range of data protection rights. Hence, vulnerable data subjects might, in fact, be subject to a lesser degree of protection.

In these cases, it is not clear how researchers should decide which is the most adequate legal basis in connection with vulnerable data subjects. Even though these subjects are well identified, the data controller would need to understand whether consent is an adequate legal basis for them. In cases where it is not, legitimate interest might be an alternative, but in such cases the data controller might be incapable of positively concluding the balancing test required in Article 6(1)(f), which concerns vulnerable data subjects.[54]

Relevance and impact on R&I

Researchers often deal with vulnerable data subjects: the ill, the frail, patients, the elderly, children, etc. These persons could be incapable of granting consent (in case of decisional vulnerability) or might be harmed during the research project (due to their physical or psychological frailty). Also, these persons could be harmed more than other 'average' data subjects in cases where these data are transferred to other data controllers for other purposes.

The responsibility of research entities to protect their data is therefore relevant, but the ambiguity of the concept of vulnerability leads to certain vulnerable subjects suffering even more disadvantages and problems.

Mitigation measures

A clearer definition of vulnerability, and list of safeguards for vulnerable subjects, is necessary. There is also a need to create guidelines for the parameters/criteria to define vulnerable

Éloïse Gennet, Roberto Andorno and Bernice Elger, 'Does the New EU Regulation on Clinical Trials Adequately Protect Vulnerable Research Participants?' (2015) 119 Health Policy (Amsterdam, Netherlands) 925. ibid.

[54] See WP29 Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC, adopted on 9 April, 2014, wp217, page 40: '[Within the balancing test], the status of the data subject is also relevant. While the balancing test should in principle be made against an average individual, specific situations should lead to a more case-by-case approach: for example, it would be relevant to consider whether the data subject is a child or otherwise belongs to a more vulnerable segment of the population requiring special protection, such as, for example, the mentally ill, asylum seekers, or the elderly'.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

individuals, and a non-exhaustive list of examples that will help researchers to identify vulnerable subjects and address their vulnerability.

1.6 Deceased people

Context and legal background

Often, researchers process the data of deceased individuals (postmortem data processing). The GDPR does not regulate postmortem data processing. Indeed, Recital 27 states that[55] "[t]his Regulation does not apply to the personal data of deceased persons. Member States may provide for rules regarding the processing of personal data of deceased persons!"

Some Member States, such as France, Italy and Spain, regulate post-mortem privacy, but many others have not. France is developing an innovative regulation regarding this type of data. According to Art. 40(1) of the French Data Protection Act, individuals can give to data controllers **general or specific indications about the retention, erasure and communication of their personal data after their death**.

In Italy, Article 2 *terdecies* of the Data Protection Act states that data subjects' rights can be exercised by those who have a personal interest, or behave to protect the deceased person, on their behalf, or for family reasons that shall be protected. This is true unless the data subject has expressly prohibited it. In any case, the data subject can withdraw their prohibition at any time. Prohibition cannot affect the economic interests of third parties, nor their right to defense. This provision has to be balanced with other rights, however. In the past, the Data Protection Authority has prohibited the daughter of a poet from accessing his email account, in order to protect the opposite fundamental right relating to the mail confidentiality of third parties.

On the other hand, in Spain, heirs have a right to access, request deletion and rectify the relevant data from data controllers and processors, unless deletion or rectification was prohibited by the deceased individual or by law. This prohibition must not affect the heirs' right to access the patrimonial data of the deceased, as pointed out in Article 3 of the Spanish Data Protection Act.

Issue

[55] See also: Answer by the Commission: **Question reference:** E-007611/2017. At: https://www.europarl.europa.eu/doceo/document/E-8-2017-007611-ASW_EN.html

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

The legal framework regarding data related to deceased people is unclear. If we consider that they are not personal data, this might mean that they can be used without consent, since the GDPR's guarantees are not applicable to these data. However, it seems unethical to use someone's data after their death without permission. Moreover, personal data from the deceased data subject, such as genetic data, could reveal personal data from their relatives and, therefore, incur a situation as described in section 1.1.5.

Also, from a legal point of view, rules on postmortem protection of honour should be taken into account,[56] since these can be applicable to some situations of postmortem data processing. As the Federal Constitutional Court of Germany stated, "[i]t would be inconsistent with the constitutional mandate of the inviolability of human dignity, which underlies all basic rights, if a person could be belittled and denigrated after his death. Accordingly, an individual's death does not put an end to the state's duty under Art 1.1 GG to protect him from assaults on his human dignity."[57] Hence, it will be necessary to take into account if postmortem data processing can imply damages to honour or any other remaining-after-death rights, regardless of postmortem data-processing rules which, for the moment, depend on the Member States' laws.

Relevance and impact on R&I

Recent words from the European Commission must be brought up here: "Data-driven innovation will bring enormous benefits for citizens, for example through improved personalized medicine, new mobility and through its contribution to the European Green Deal"[58]. Nonetheless, as the Commission states, "citizens will trust and embrace data-driven innovations only if they are confident that any personal data sharing in the EU will be subject to full compliance with the EU´s strict data protection rules".

[56] In Spain, a wide and recent study on the issue has been developed by Minero Alejandre, G. La protección postmortem de los derechos al honor, intimidad y propia imagen y la tutela frente al uso de datos de carácter personal tras el fallecimiento. Aranzadi. 2019.
[57] Federal Constitutional Court (First Senate) 24 of February 1971. BVerfGE 30, 173. Available at: https://germanlawarchive.iuscomp.org/?p=56.
[58] European Commission. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – A European strategy for data. Brussels. 19.2.2020 COM (2020) 66 final. p. 2. Available at: https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

But what happens if the data that can be useful to, for instance, personalized medicine researchers, are data from dead people? Are any EU applicable rules? Is there any guidance on this topic while the EU or the Member States develop rules on it?

If we consider that the GDPR no longer applies to personal data of deceased persons, they can be used for research freely. This would open up an impressively extensive database for research purposes. The researcher could, perfectly legally, make use of these data (maybe with the only limitation expressed by the need to respect the honor of the deceased person) and this might be problematic as it could also reveal data (genetic data, illnesses, etc.) of living relatives (descendants, brothers, sisters).

An alternative interpretation of the GDPR might be that its provisions are applicable to data that were personal when collected by the controller. Nonetheless, there remains many questions about this. In the public domain when the person passes away, or should a legal regulation similar to postmortem organ donation be developed?[59] Which specific issues should be included in this future regulation?

This topic is deeply related to testamentary freedom and postmortem privacy, which has been a new research topic in recent years.[60] The importance of this issue is even greater when considering data pertaining to a special category: genetic data. These data include a considerable amount of sensitive information about the relatives of deceased people. If we consider that they are no longer personal data, we would be performing research with data that reveals information about living citizens' health - without their consent or even their knowledge.

Mitigation measures

A common European code of conduct for researchers should suggest common rules for processing the personal data of deceased data subjects, in order to reach a better harmonization

[59] The option of developing a legal regime similar to portmortem organ donation has been defended by Krutzinna, J., Taddeo, M. & Floridi, L. Enabling Posthumous Medical Data Donation: A Plea for the Ethical Utilisation of Personal Health Data. Sci Eng Ethics. 2019 Oct. 25 (5): 1357-1387.

[60] See Harbinja, E. Post-mortem privacy 2.0: theory, law and technology. International Review of Law, Computers & Technology. 31 (1). 2017. 26-42. The author is in favour of translating freedom of testation into the online environment, where digital assets mainly compromise of informational and personal data content since individuals should be able to exercise their autonomy online and decide what happens to their assets and privacy on death.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

and avoid uncertainty for international research projects. In the case of genetic data, a wide interpretation of the concept of personal data would make it possible to give relatives the chance to consent or not to data processing for research purposes. This would impede the use of data related to their health without their consent.

While other consensuses are reached in the EU, another mitigation measure could be developing, as part of this common European code of conduct for researchers: a formula to enable people to give consent to their personal data after death for research purposes, when provisions of GDPR on this topic are not enough (e.g. genetic illnesses research). If considered necessary, this formula would include consent of relatives when needed.

1.7 Scope of the DPIA and the notion of data subjects

Context and legal background

The scope of the DPIA (Article 35 of the GDPR) refers to the processing of personal data that has a high risk for the fundamental rights and freedoms of the data subjects. However, Article 35(1) names the "fundamental rights and freedoms of *natural persons"'* (emphasis added), not just of data subjects. In addition, Article 35(7) (d) states that data controllers should identify "the measures envisaged to address the risks (…) taking into account the rights and legitimate interests *of data subjects and other persons concerned*" (emphasis added).

Issue

It is not clear who these 'other persons concerned' might be, or whether and how data controllers could take those individuals into account. It could refer to the general group that may be impacted by some data-processing operations, or to other individuals whose data might be inferred from processing the data of the subject directly involved in the research (e.g. relatives).

Relevance and impact on R&I

In research, impacted individuals are not the only data subjects. As discussed earlier, there are other data subjects (see Section 1.4.5) that could be affected by a data-processing activity. Taking a wider perspective, all research could have an impact on other individuals, for example through its results. Therefore, there is high uncertainty about whether a broader definition of

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

impacted individuals should be considered in the risk assessment within the DPIA of research projects, and which other data subjects should be taken into account.

Mitigation measures

Such uncertainty should be clarified through interpretative guidelines that address the issue of impacted individuals and provide criteria to determine the groups of subjects that should be taken into account.

1.8 Safeguards for research purposes in Article 89

Context and legal background

The GDPR is designed to protect data subjects but also allows for data processing in scientific research. Article 89(1) states that the possible use of data processing "for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimization. Those measures may include pseudonymization provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner."

Issue

The vagueness in the terms on the list of desirable safeguards is a problem (e.g. for small research entities or SMEs that process personal data for the first time). Only 'pseudonymization' and subsequent anonymization of data, in accordance with the data minimization and storage limitation principles, are mentioned as examples. In terms of transparency, for example, there is no clear guidance in the body of Article 89.[61] This is problematic, because Article 89 is the only safeguard in many cases, such as data processing

[61] See, however, the proposal of transparency-by-design in research in Arianna Rossi and Gabriele Lenzini, 'Transparency by Design in Data-Informed Research: A Collection of Information Design Patterns' (2020) 37 Computer Law & Security Review 105402.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

for secondary purposes. As stated in Section 1.4.2, Article 5(1)(b) allows for further data processing purposes in case of research purposes, without the need to perform a compatibility test under Article 6(4). The lack of a compatibility test should be counterbalanced by appropriate transparency and accountability measures to better protect the data subject.

Relevance and impact on R&I

This vagueness could lead to a minimalist approach when implementing Article 89. For example, researchers could implement only data pseudonymization or subsequent anonymization, and ignore other safeguards.

To counter this, WP29 has indicated that transparency could constitute a more than appropriate measure to ensure compliance with Article 89: "Transparency is an additional safeguard when the circumstances of the research do not allow for a specific consent."[62] Moreover, the proper implementation of the principles prescribed in Article 5 could provide additional safeguards. For example, data minimization could contribute to researchers collecting just the necessary information, or limiting the storage of data once the purposes of the research have been achieved instead of keeping the data for further research (although this is possible, as discussed above). Researchers could adopt all steps necessary to validate the accuracy of the information to avoid collecting false or incorrect information, for example by only collecting data from the data subjects directly.

Also, according to Article 89, Member States' legislations could adopt specific safeguards. However, this means that the application of safeguards for data processing for research purposes might differ among Member States, which could create problems in particular for research projects based in more than one Member State (see Section 1.3).

Mitigation measures

There is no clear list of safeguards for data processing for research purposes according to Article 89. This could be created through a list of safeguards, included in a European code of conduct for researchers, or in interpretative documents from the EDPB. This could also include specifically tailored DPIAs, more precise information duties for data controllers, and more limits to collecting the data of vulnerable data subjects, among other contents.

'Guidelines on Consent under Regulation 2016/679' (n 9) 29.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

Data anonymization should be the most encouraged and promoted safeguard, following the wording prescribed by Article 89(1), which states that "[w]here those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner". WP29 highlighted this measure in the following manner: "Anonymisation is the preferred solution as soon as the purpose of the research can be achieved without the processing of personal data."[63] Therefore, the implementation of anonymization measures to remove a compliance burden from research entities, and allow them to conduct their intended research, should be encouraged when possible.

Moreover, specific practices and tools should be suggested to researchers to properly self-enforce the principles prescribed in Article 5, and to add safeguards to their data-processing activities, with particular attention given to transparency, storage limitation and data minimization.

1.9 Special categories of personal data

Context and legal background

As a general rule, the GDPR does not allow for the processing of personal data that fall under certain special categories - unless the data controller can rely on one of the exemptions prescribed in Article 9(2). These categories are defined as follows in Article 9(1): "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation".

Issue

While researchers might be able to demonstrate that they can rely on the exemptions to process special categories of personal data, they might have issues determining whether the data collected actually belongs to a special category of personal data. Its definition is far from settled,

ibid.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

given its broad potential reach, as the wording of Article 9(1) prescribes. Therefore, the boundary between personal data and special categories of personal data is often blurred.[64]

One special category of personal data that is commonly involved in research is health data. Recital 35 states that "personal data concerning health should include all data pertaining to the health status of a data subject which reveal information relating to the past, current or future physical or mental health status of the data subject". Consequently, some personal data (that cannot be considered directly as a special category of personal data) might allow the research entity to infer special categories of personal data.

For example, lifestyle data, (e.g. data related to daily diet or fitness) could be considered as included in such a definition, as it can relate to the data subject's health.[65] Following this example, the collection of information related to sugar or wheat consumption under research on eating habits could indicate which data subjects might be diabetic or coeliac, respectively, and therefore the research entity might be collecting special categories of personal data that it did not intend to collect originally. Similarly, location data reveal many potentially special categories of personal data (e.g. sexual data, sexual orientation, political beliefs, religious beliefs), but it is not clear whether such data should be always considered as a special category of personal data.

Relevance and impact on R&I

Researchers often process a large amount of data that could be considered as belonging to special categories of personal data (e.g. medical data, lifestyle data, location data). If such data are processed on a large scale, this could oblige researchers among other steps, to have a data protection officer (DPO)[66], according to Article 37, and to perform a DPIA according to Article 35(1). These accountability duties would require financial and organizational efforts: understanding whether the processed data are included in a special category of personal data or not is therefore extremely relevant for researchers.

[64] See Malgieri and Comandé (n 21); K Mccullagh, 'Data Sensitivity: Proposals for Resolving the Conundrum' (2007) 2 Journal of International Commercial Law and Technology 190.
[65] Giovanni Comandè and Giulia Schneider, 'Regulatory Challenges of Data Mining Practices: The Case of the Never-Ending Lifecycles of "Health Data"' (2018) 25 European Journal of Health Law 284.

[66] Public research entities are already obliged to have a DPO.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.

Mitigation measures

While the EDPB might release some interpretative guidelines on the notion of special categories of personal data, and clearer criteria (and examples) on how to delimit the border between what is a special category of personal data and what is not, these documents cannot provide an exhaustive list, given the broad definition of both personal data and special categories of personal data. Therefore, alongside the issuance of guidance on this matter, the EDPB should produce regular summaries with the case law from local DPAs, where the concept of special categories of personal data is interpreted. This will allow data controllers (in this case research entities) to develop a broad and complete picture of the special categories of personal data found across the EU.

As an alternative, the EDPB could clarify at which level some general criteria (e.g. a contextual criterion or purpose-based criterion[67]) should be adopted to interpret the categories of personal data in GDPR Article 9(1).

[67] Schwartz, P.M. & Solove, D.J. (2014). Reconciling personal information in the United States and European Union. *California Law Review*, 102 (2014), 877.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document is the property of the PANELFIT consortium and shall not be distributed or reproduced without the formal approval of the PANELFIT Project Coordinator.