COMPUTER LAW & SECURITY REVIEW 41 (2021) 105539

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/CLSR

Computer Law & Security Review

Comment

Can the GDPR make data flow for research easier? Yes it can, by differentiating! A careful reading of the GDPR shows how EU data protection law leaves open some significant flexibilities for data protection-sound research activities ☆

Giovanni Comandè[a,1], *Giulia Schneider*[b,1,*]

[a] *Full Professor of Private Comparative Law at Sant'Anna School of Advanced Studies, Pisa, Italy*
[b] *Postdoc Researcher in Private Comparative Law at Sant'Anna School of Advanced Studies, Pisa, Italy*

ARTICLE INFO

Keywords:
Health data
Research
Data protection
Free flow
GDPR
Safeguards

ABSTRACT

Against the common perception of data protection as a road-block, we demonstrate that the GDPR can work as a research enabler. This study demonstrates that European data protection law's regulatory pillars, the first related to the protection of the fundamental right to data protection and the second regarding the promotion of the free flow of personal data, result into an architecture of layered data protection regimes, which come to tighten or relax data subjects' rights and data protection safeguards *vis à vis* processing activities differently grounded in public or merely economic interests. Each of the identified data protection regimes shape different "enabling regulatory spots" for the processing of sensitive personal data for research purposes.

☆ This paper has been developed within the "SoBigData Plus Plus: European Integrated Infrastructure for Social Mining and Big Data Analytics" Project, funded by the EU Commission under the H2020 INFRAIA-1-2019 program (GA 871042) and within the "Legality Attentive Data Scientists (LeADS)" Project, funded by the EU Commission under the H2020-EU.1.3.1. - Fostering new skills by means of excellent initial training of researchers (GA 956562)

* Corresponding author: Giulia Schneider, Sant'Anna School of Advanced Studies, Pisa, Italy.
E-mail addresses: giovanni.comande@santannapisa.it (G. Comandè), giulia.schneider@santannapisa.it (G. Schneider).

[1] These authors contributed equally to this work, however paras 1–3 are to be attributed to Giulia Schneider, whereas paras 4–6 to Giovanni Comandè.

https://doi.org/10.1016/j.clsr.2021.105539
0267-3649

1. Today's data-driven research and the GDPR

The definition of the boundaries of openness of datasets allowed by the GDPR in scientific research is an exercise of particular importance because of the prevalence of data protection law in 'inextricably linked' datasets[2] and of even greater relevance for cross-border projects, orphans also of the Privacy Shield.[3]

In this context, the GDPR framework for data-driven research is object of a twofold line of critiques, swinging from those regarding the GDPR as a direct impairment to research,[4] and those blaming the lack of clear interpretative guidelines for research as the actual brake on scientific progress.[5]

Here we demonstrate the contrary: the GDPR is the cornerstone around which a safe and free data flow for research data can be built. This explains the deference given by all other data-related legislations. We will illustrate how the GDPR fosters a reliable and balanced framework for data sharing and related research objectives, to the very opposite as what the mainstream literature pictures. Here we show the existence of "differential" data protection regimes for research, resulting from the grading of allowed exceptions and required safeguards- thus of related standards of protection-, in consideration of the different nature of the research at stake.

2. The GDPR's frameworks for research

While data pooling is being encouraged by European regulators,[6] players are still searching ways to overcome the perceived major roadblock to the aggregation of research valuable data in the European data space: data protection law. It is not clear to many which is the GDPR's[7] stand in respect to legislations aiming to open up businesses' and public entities' datasets for research purposes.[8] These legislations declare they do not "affect the protection of individuals with regard to the processing of personal data under Union or national law".[9] Thus, they defer the regulation of data sharing to the GDPR.

Research has a particularly important role within the GDPR. Recital 157 GDPR declares the GDPR's aim of "facilitating" research, an aim reflected by the many legitimate bases offered for the processing of special categories of data in research, as consent (art. 9(2)a GDPR); the need or protection against serious cross-border threats to health or the safeguards of high standards of quality and safety of health care and of medicinal products or medical devices (art. 9(2)i GDPR); the conduction of research activities (art. 9(2)j GDPR).

These legal bases can be combined with the specific rules the GDPR sets for research (so called research exceptions). Amongst them, there are the default compatibility with the purpose limitation principle on further processing for research purposes under arts. 5(1) lett. b) and 6(4) GDPR; the

[2] Art. 2(2) Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union, OJ L 303/59 (28 November 2018) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807&from=EN. European Commission, Communication from the Commission to the European Parliament and the Council – Guidance on the Regulation on a framework for the free flow of non-personal data, COM(2019) 250 final (29 May 2019) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019DC0250&from=EN.

[3] Court of Justice of the European Union C-311/18 Data Protection Commissioner v. Facebook Ireland Limited, Maximilian Schrems (16 July 2020) http://curia.europa.eu/juris/document/document.jsf?text=&docid=228677&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=9745404. European Data Protection Board, 'Recommendations 1/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data' (10 November 2020) https://edpb.europa.eu/sites/edpb/files/consultation/edpb_recommendations_202001_supplementarymeasurestransferstools_en.pdf.

[4] See David Peloquin & Michael Di Maio & Barbara Bierer & Mark Barnes, 'Disruptive and Avoidable: GDPR Challenges to Secondary Research Uses of Data' (2020) 28 European Journal of Human Genetics 697–705; Tania Rabesandratana, 'European Data Law is Impeding Studies on Diabetes and Alzheimer's, Researchers Warn' (20 November 2019) Science https://www.sciencemag.org/news/2019/11/european-data-law-impeding-studies-diabetes-and-alzheimer-s-researchers-warn; Birgit A. Simell et al., 'Transnational Access to Large Prospective Cohorts in Europe: Current Trends and Unmet Needs' (25 March 2019) 49 New Biotechnology 98–103; Andreas Wiebe & Nils Dietrich, *Open Data Protection: Study on Legal Barriers to Open Data Sharing – Data Protection and PSI* (Universitätsverlag Göttingen, 2017) https://univerlag.uni-goettingen.de/handle/3/isbn-978-3-86395-334-8; Lothar Determan, 'Healthy Data Protection Law' (2020) 26, 2 Michigan Technology Law Review 229–278.

[5] See Robert Eiss, 'Confusion over Data-privacy Law Stalls Scientific Progress' (27 August 2020) 584 Nature 498.

[6] European Commission, Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act), COM(2020) 767 final (25 November 2020) https://ec.europa.eu/digital-single-market/en/news/proposal-regulation-european-data-governance-data-governance-act, at recital 35. 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' (19 February 2020) COM(2020) 66 final https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN.

[7] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119/1 (4 May 2016) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN, hereafter GDPR.

[8] Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union, OJ L 303/59 (28 November 2018) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807&from=EN. Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending directives 96/6/EC and 2001/29/EC, OJ L 130/92 (17 May 2019) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L0790&from=EN. Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-use of Public Sector Information, OJ L 172/56 (26 June 2019) https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L1024&from=EN.

[9] Recital 52 Directive (EU) 2019/1024 (n 5) and recital 8 Regulation (EU) 2018/1807.

derogation to other general principles as the storage limitation principle under 5(1) lett. e) GDPR and to some data subjects' rights, as the right to erasure under art. 17(3) lett. c) and e) GDPR and the right to access under art. 14(5) lett. b) GDPR; finally, art. 89(1) GDPR requiring controllers to enact appropriate measures to safeguard data subjects' fundamental rights and freedoms that may be impaired in the course of research investigations.

To give content to accountability and limit abuses by data controllers of the research exceptions the "essence of the (fundamental) right to data protection" must be respected.[10] Following such a cautious interpretative approach, a possible restraint to creeping abusive applications of the research exceptions can be found in the distinction between public interest and commercial-oriented research, without penalizing commercial research as such.

Against this backdrop, we argue that the GDPR lays down a multilayered framework for data processing for research purposes, sensitive to the types of research pursued.

3. 'Differential' research in the GDPR

Recital 159 GDPR sustains the extension of the research exceptions to private motivated/funded research. This wide notion of research is confirmed also by the recently issued proposal for a Data Governance Act.[11], in lin , in line with the European Commission's Strategy for data, which has stressed the public good dimension of data shared among both private and public entities.[12] The GDPR thus leads an approach different from the one taken by most US data protection regulations, as the CCPA. The CCPA, indeed, narrowly defines and circumscribes the notion of research only to public-interest oriented research activities, excluding from the scope of research-oriented processing activities those serving commercial purposes.[13]

However, there are indications in the GDPR that distinguish among research with advantages to data subjects and the general public, and research mostly profit-oriented.

The same recital 159 GDPR signals a possible differentiation of data protection regimes echoed in the need to take into consideration "reasons for further measures in the interest of the data subject". If we assume as an example research for orphan diseases, it clearly "gives reason for further measures in the interest of the data subject". Thus, the fact that "the general rules of this Regulation should apply in view of those measures" shows in turn that the GDPR regime could be softened in case of public interest-oriented research processing activities.

Hence, the differential regimes for research, while do not differentiate among private and public funding, clearly differentiate in terms of the more "egoistic" or "altruistic" aim of the research. The subjective perspective regarding the private or public nature of the funding, and thus the private or public nature of the entities conducting research, appears to be quite irrelevant since it can well be the case that also privately-funded research serves broader public interest goals, as it can occur with the research and development of a vaccine during a pandemic.[14]

Conversely, the objective perspective highlighting that the public interest and commercial-based research activities are linked is very informative. In this respect, as recital 159 GDPR seems to indicate, the boundary is to be drawn between research whose results benefit also data subjects and research which, as acknowledged under the CCPA, mainly serve controllers' economic interests. However, differently from the CCPA, the "broad interpretation" of the notion of research recalled by the same recital 159 GDPR suggests including in the research-based data protection regimes also the latter type of research.

Against this backdrop, we propose to employ such distinction for the purposes of scaling the flexibilities or "privileges" – as the German Data Ethics Commission defines them[15]– granted by the GDPR to processing activities conducted for research purposes.

4. 'Differential' data protection regimes for different research

The different interaction between the legal bases for personal data processing for research and the recalled data protection framework for research creates a dynamic spectrum of legal regimes ranging from data subjects' full control (consent) for private data pools processed for-profit purposes to data subjects' transfer of control to data controllers for private or public data pools employed for-non-profit/public interest research-oriented purposes. The research exception is actually plural!

In this perspective, we identify a 'data subject-based'; a 'public interest-based' and a purely 'research-based regime'. Under the first regime the data subjects' rights provided by Chapter III GDPR are fully actionable; under the public

[10] European Data Protection Supervisor, A Preliminary Opinion on Data Protection and Scientific Research (6 January 2020) https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf, 18.

[11] European Commission, Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act), (n 5) at recital 35.

[12] 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', (n 5) 6–7.

[13] See section 1798. 140 lett.d) and f) CCPA. For the literature see William Nicholson Price & Margot E. Kaminski & Timo Minssen & Kayte Spector-Bagdady, 'Shadow Health Records Meet New Privacy Laws- How Will Research Respond to a Changing Regulatory Space?' (2019) 363, 6425 Science 448, 450.

[14] European Data Protection Board, Guidelines 03/2020 on the Processing of Data Concerning Health for the Purposes of Scientific Research in the Context of the Covid-19 Outbreak (21 April 2020) https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202003_healthdatascientificresearchcovid19_en.pdf, para 64.

[15] Bundesministerium für Justiz und Verbraucherschutz, 'Opinion of the Data Ethics Commission' (22 January 2020) https://www.bmjv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN_lang.html;jsessionid=776E9EC21856B458D8F4D1927D7705C5.1_cid324?nn=11678512, 124.

interest-based regime some derogations to ordinary data subjects' rights may be established by Union or national laws in accordance with art. 23 GDPR;[16] conversely, under the research-based regime, substantial derogations to those rights are envisaged directly in the GDPR and further ones can be introduced by State and Union law in accordance with art. 89(2) GDPR. However, in order to counterbalance the weakening of actionable data subjects' rights, the GDPR shifts the burden of care onto data controllers, which are required under art. 89(1) GDPR to enact adequate safeguards for the protection of data subjects' rights and freedoms: with greater powers come greater responsibilities managed under the accountability principle.

5. Shaping the GDPR's differential data protection regimes for research

Even under the restrictive approach required by the EDPS,[17] it is possible to differently modulate the GDPR's flexibilities for public interest-oriented (or altruistic) research and profit-driven one, regardless of the sources of their funding. Such modulation is primarily rooted in the principles of proportionality and fairness, protecting data subjects from controllers'/processors' abuses.

In this perspective, data subjects' control rationales and free flow of information goals are the parameters upon which the taxonomy is based. Control rationales suggest that both the derogations and the safeguards required under the research-based regimes should be respectively restricted to the minimum and stretched to the highest when it comes to merely commercially-oriented research data processing. Conversely, public interest-oriented research activities could enjoy a more enabling regulatory regime, designed around deeper derogations, if needed, defined at national level and less burdensome safeguards, facilitating the flows of research data.

5.1. *From public interest to for profit research*

In the Italian Tiziana case[18] many citizens volunteered pursuing data-philanthropy aims that leaded the transfer from a public good oriented institution to a for-profit one to fail the further use test today envisaged by art. 6(4) GDPR.[19]

The principles of proportionality and necessity would require processing activities conducted for for-profit research purposes to rely on the legitimate basis that is more respectful of data subjects' interests and rights: consent and the related possibility of its withdrawal structurally assures a higher degree of control, also if it is related only to broad research areas as suggested by recital 33 GDPR (e.g. for data philanthropy as in the Tiziana case). The effectiveness of data subject's control under consent is mitigated by the presumption of compatibility under arts. 6(4) and 5(1) lett. b) GDPR, enabling further processing for research purposes. Yet, the same mentioned principles of proportionality and necessity impose a strict interpretation of the compatibility rules restricting further flows of research data to data subjects' self-informational determinations (e.g. through notice requirements).

Hence, changing the context (art. 6(4)(f) GDPR) from merely altruistic goals to also for-profit ones might lead to failing the compatibility test. Equally, the derogations to ordinary data protection rights could be circumscribed to the sole derogations directly allowed by the Regulation and not be aggravated by Member States laws.

Conversely, also for-profit research can benefit of the derogations as long as appropriate safeguards are provided, for instance selecting "processing which does not require identification" (art. 11 GDPR). This possibility is directly drawn from article 89(4) GDPR establishing a principle of segregation: privileges only apply to research purposes and do not extend to other purposes. A striking example can be offered by research for marketing and the use of the research outputs for marketing. While personal data processing for scientific studies on marketing would enjoy the research privileges, the use of the same data for purely marketing purposes would not, in consistency with the principle of segregation illustrated by recital 162 GDPR stressing that the results of statistical purposes processing operations should not be used "in support of measures or decisions regarding any particular natural person".

On the side of the safeguards required under art. 89(1) GDPR, the principles of proportionality, fairness and segregation suggest the enactment of higher context-sensitive safeguards for preventing research processing activities to result into the processing of data for pure (*i.e.* not research-based) commercial purposes.

5.2. *Mixing interests in public-private research*

In the case of mixed private-public health datasets employed for research purposes, the data protection research regime can be calibrated based on the influence that commercial undertakings have within established research partnerships or organisations. The degree of influence of these entities, indeed,

[16] See art. 23 (1) lett. e) GDPR specifically referring to public health concerns.

[17] European Data Protection Supervisor, A Preliminary Opinion on Data Protection and Scientific Research (6 January 2020) https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf, 18.

[18] Tribunal of Cagliari, Sentenza n. 1569, 6 June 2017.

[19] The Italian Data Protection Authority has blocked the transfer of genetic data between the Italian genomic biobank Shard.Na and UK-based for-profit corporation Tiziana Life Science plc. So Italian Data Protection Authority, Provvedimento di blocco del trattamento dei dati personali contenuti in una biobanca n. 389 (6 October 2016) https://www.garanteprivacy.it/pdf?p_p_id=PdfUtil&p_p_lifecycle=2&p_p_state=normal&p_p_mode=view&p_p_resource_id=%2Foffering%2FprintPDF&p_p_cacheability=cacheLevelPage&_PdfUtil_articleId=5508051. Tribunal of Cagliari, n. 1569, 6 June 2017, overturned the decision. A subsequent decision by the Italian Data Protection Authority, again ordered the English company to block the processing of health data of those data subjects that had withdrawn their consent. Italian Data Protection Authority, Provvedimento 21 dicembre 2017 n. 561 (21 December 2017) https://www.garanteprivacy.it/pdf?p_p_id=PdfUtil&p_p_lifecycle=2&p_p_state=normal&p_p_mode=view&p_p_resource_id=%2Foffering%2FprintPDF&p_p_cacheability=cacheLevelPage&_PdfUtil_articleId=7465896.

determines the risk of commercial "capture" of research results (when for-profit interests weight in).

The involvement of for-profit organisations and thus their influence in the governance of research projects and results can be derived from specific parameters. As stated, under the GDPR it is not who funds the research that matters, but its scope.

In the collaboration between private and public actors, as in the "Innovative Medicines Initiative",[20] based on a public-private partnership between the European Commission and the pharmaceutical industry, the mere presence of commercially-oriented stakeholders might trigger the enactment of higher data protection safeguards and lower derogations from the ordinary regime. Nonetheless, purposes of public health protection, and the need of immediate research actions, suggest a relaxation of data protection checkpoints. In the specific cases where mixed health data pools are employed for research purposes in the public interest in the area of public health, such as for the protection against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices, the higher level of restrictions on the processing of special categories of personal data can be eased, in accordance with what is required for the processing for public interest purposes under art. 9(2) i) GDPR, disregarding the public or private nature of the subjects involved.[21]

6. Conclusions

Differential data protection regimes for research are rooted in the GDPR's double fine-tuning system based on the balancing among coded data protection principles and rules and the establishment by data controllers of adequate safeguards for the protection of data subjects' rights and freedoms.

We read in the GDPR a scaling of this double fine-tuning system in respect to different research-based processing activities over sensitive data. This means that in case of merely for-profit research activities data subjects are entitled to a greater control over occurring processing operations due to a fuller application of data protection principles and rights and a more severe layer of safeguards that controllers need to enact; conversely, in case of public interest-based research possible derogations can be exploited with greater ease by data controllers, which can establish lighter additional safeguards. Such differential data protection regimes have thus a varied, and 'differential' impact on the contractual freedom to share and aggregate personal data, also of sensitive nature.

From this angle, the GDPR provides a highly sophisticated regulation of data processing activities for research purposes: in the aim of complementing sharing objectives with the high level of protection for data subjects' fundamental rights, it balances research privileges and individual rights privileges, by variously scaling them in respect to the nature of conducted research activities.

The sensitivity of the European data protection regulatory model could inspire also the developments of US data protection regulations for research purposes since it offers a pro-research set of differential regimes able to foster data flows without hampering the essence of the fundamental right to personal data protection. A much-needed solution after the final collapse of the Privacy Shield.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

No data was used for the research described in the article.

[20] See IMI-Innovative Medicines Initiative, https://www.imi.europa.eu/.

[21] This interpretation is suggested also by Mike Hintze, 'Science and Privacy: Data Privacy Laws and their Impact on Research' (2019) 14 Washington Journal of Law, Technology & Arts 103, 134.