# Scientific research processing health data in the European Union: data protection regime vs. open data

Giorgia Bincoletto*
*Post-Doc Research Fellow at the Faculty of Law, University of Trento

**Abstract.** Data processing is fundamental for medical and biomedical scientific research. Data access and open sharing foster innovation and knowledge in these fields. However, supporting data-intensive research raises several data protection issues. The legal framework is complex and the proposal recently issued by the European Commission on the European Health Data Space aims at introducing rules to open research data. This paper provides an overview of the relevant data protection laws in the European Union, focusing on the key elements for researchers. It also deals with the policies and legislation on Open Science and Open Data, and discusses the new acts within the Data Strategy that may be relevant for a scientific project to open medical research. Then, the tensions between the data protection regime and the open research data framework are analysed to find the viable solutions for data access and sharing.

**Keywords:** health data, scientific research, data protection, open data, EHDS

## 1. Introductory remarks

Medicine and biomedical sciences are data-intensive fields. It may be argued that data processing is essential for the research carried out in these domains: clinical trials typically collect and aggregate participants' data to achieve their scientific outcomes[1], biobanking research combines biological samples with health data (Hallinan, 2021) (Slokenberga et al., 2021), prospective and retrospective studies also require such information and, ultimately, all projects aim at publishing and sharing their results. Scientific research is necessary for the progress of the healthcare sector that is increasingly focused on digital technologies.

Innovation and knowledge depend on access and data sharing. It has been stated that enabling researchers' access to large volume of health data accelerates improvements that benefit both clinical practice and public health (Shabani et al., 2021, 187). The availability of data and research materials is also necessary to the reproducibility and replica-

[1] Clinical trials are usually listed at ClinicalTrial.gov that is a resource provided by the U.S. National Library of Medicine, but it is also used around the world to advertise privately and publicly funded studies.

bility of the experiments[2]. However, supporting data-intensive research raises several legal issues.

Focusing on the European Union (EU) context, the lawfulness of collecting and using health data for research purposes needs to be addressed from multiple perspectives. On the one hand, data should be protected according to data protection law when the initiatives process "personal health data", meaning information on the past, present and future physical or mental health status of an identified or identifiable natural person[3]; on the other hand, the Open Science, Open Access and Open Data policies aim at opening up scientific research, including the collected information and outcomes (Margoni, Caso, Ducato, Guarda, 2016). The goal of these initiatives to ensure that data are "as open as possible, as closed as necessary" has been included in some legal provisions[4]. Openness implies making data available to anyone without financial, technical and less legal constraints. Data protection could occasionally lead, on the contrary, towards closure to safeguard fundamental rights, but amongst its principles it is also stressed that the free movement of personal data within the EU should be neither restricted nor prohibited[5].

Data protection is extremely important to safeguard the data subject's individual rights, especially in the medical sector where intimacy is essential (Hansson, 2021, 34) (Madir, 2020). Medical confidentiality is indeed a general principle in the healthcare domain and the law provides a duty of keeping information private (Véliz, 2019). The Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects prescribes that privacy of research subjects should be always protected[6]. At the same time, access and sharing of health data are indispensable for research, and under the Open Data paradigm they advance the public interests. Some authors even promote a "human right to Open Science" (Caso, Binda, 2020). In the "European Strategy for Data" of 2020 the European Commission argued that the use and re-use of health data could benefit the healthcare sector[7].

[2] *See* the report of the National Academies of Sciences, Engineering, and Medicine of 2019 "Reproducibility and Replicability in Science", available at ww.ncbi.nlm.nih.gov/books/NBK547546/.

[3] On the definition *see* further Section 2.

[4] *See* further Section 3.

[5] *See* further Section 2.

[6] *See* paragraph 24 of the Declaration available at https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/.

[7] The European Commission stated that "strengthening and extending the use and re-use of health data is critical for innovation in the healthcare sector. It also helps healthcare authorities to take evidence-based decisions to improve the

The recent pandemic has shown the urgent need to share scientific progress and accelerate innovation (Besançon et al., 2021) (Caso, 2020). The legal framework is complex, and the legislative proposals recently issued by the European Commission seek to introduce new rules to foster the open movement of data that may be applicable in the healthcare sector. The creation of a European Health Data Space is the primary example of this trend[8].

Researchers should therefore strike the balance between personal data protection and open data sharing. During the research data management has become a hard task, and it is not clear when it is lawful to open the collected information. Against this backdrop, the paper proceeds as follows: Section 2 provides an overview of the relevant data protection laws in the EU, focusing on the key elements that researchers should carefully take into account in the design and execution of their activities, and the specific provisions of the General Data Protection Regulation (GDPR) dedicated to scientific research. Section 3 explores the open data paradigm and discusses the new legislative initiatives that may be relevant for a scientific project to make health data open. The analysis of the interplay between the data protection regime and the open data framework is provided in Section 4. Finally, Section 5 concludes with some final remarks.

## 2. The data protection framework for research studies processing personal health data in the European Union

In the EU legal framework, the general requirements on the processing of personal data are provided by the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation or "GDPR"), which entered into force on 24 may 2016 and applies since 25 May 2018[9] (Kuner, Bygrave, Docksey, 2020). Specific provisions regarding processing of personal data

accessibility, effectiveness and sustainability of the healthcare systems. It also contributes to the competitiveness of the EU's industry. Better access to health data can significantly support the work of regulatory bodies in the healthcare system, the assessment of medical products and demonstration of their safety and efficacy". *See* the "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', COM (2020) 66 final (February 19, 2020)', paragraph 4.

[8] *See* further Section 3.

[9] 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, and repealing Directive 95/46/EC (General Data Protection Regulation), O.J. L. 119, 4.5.2016.

connected with criminal offences or the execution of criminal penalties are detailed in the Directive (EU) 2016/680, which has been implemented at Member States' level, and Regulation 2018/1725 sets forth the rules applicable to the processing of personal data by European Union institutions, bodies, offices and agencies. Moreover, Directive 2002/58/EC contains special rules for the processing of personal data and the protection of privacy in the electronic communications sector (e-Privacy Directive)[10]. All these laws recognise the individual right to data protection that is guaranteed by the EU Charter of Fundamental Rights (Art. 8).

Among these frameworks, scientific research processing health data might be subject to the GDPR, even when the studies are conducted in the public interest in the area of public health[11]. The requirements are applicable both to the use and the disclosure of personal data for scientific research purposes[12].

The key data protection elements the researchers should carefully take into account during the design and execution phases and for data management and governance may be summarised as follows: 1) the application or not of the GDPR; 2) pseudonymisation and anonymisation of personal data; 3) purpose limitation and primary and secondary uses of data; 4) data minimisation and storage limitation; 5) the presence of data concerning health as a special category of data and the related legal grounds to process such data at the EU and national level; 6) data protection by design and by default.

Only the research activity that processes "information relating to an identified or identifiable natural person" (Article 4(1) GDPR) and that is carried out by a data controller according to the material and territorial scope of the GDPR (Artt. 2 and 3 GDPR) should comply with the general rules. Pseudonymised data is included in the notion of personal data since it can be "attributed to a natural person by the use of additional information" that has been separated by the data controller[13]. The researcher should define in advance whether the

[10] The Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications), COM/2017/010 final – 2017/03 (COD), is still under discussion at the EU Parliament and Council.

[11] *See* Recital 159 of the GDPR.

[12] Ibidem.

[13] Pseudonymisation process de-associates the identity of the data subject from the information processed by the use of a pseudonym. On the pseudonymisation techniques see the report of the European Union Agency for Network & Information Security, "Recommendations on shaping technology according to GDPR provision. An overview on data pseudonymisation" (2018).

project includes the use of personal data, and then it should follow the requirements of the GDPR. The data flows and the technical and organisational details of the processing activities should be carefully mapped.

When data is not personal, 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" applies in place of the GDPR[14]. A third category is not conceived by the system (Comandè, 2022, 41). Data is not personal either if it is "anonymous information, namely information which does not relate to an identified or identifiable natural person" *ab origine* or it is "personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable", meaning personal data that is later subjected to anonymisation[15] (Purtova, 2018). The researcher may use anonymous data that are already available in an open dataset or may anonymise personal data. In this last case, data collection should follow the GDPR, but after the anonymisation process the rules would not apply. However, the research goals may not be achieved with the use of non personal data.

In addition, anonymisation is a hard task because it should be effective and the re-identification should be impossible[16] (Stalla-Bourdillon and Knight, 2017). Determining whether the dataset is anonymised is complex, since datasets are often and easily de-anonymised by the use of re-identification techniques (Narayanan and Shmatikov, 2008). It has been argued that combining allegedly anonymised datasets with other outside information available online (e.g. on social networking websites) frequently pries out obscured identities. Consequently, people lose faith in anonymisation (Ohm, 2010). When it is possible to turn anonymised data into personal data and every time in a mixed dataset non personal and personal data are inextricably linked, such data are to be treated as personal data, and the GDPR applies[17].

[14] 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, PE/53/2018/REV/1, O.J. L. 303, 28.11.2018.

[15] *See* the working of Recital 26 of the GDPR. Examples of non personal data are listed in Recital 9 of the Regulation 2018/1807: "aggregate and anonymised datasets used for big data analytics, data on precision farming that can help to monitor and optimise the use of pesticides and water, or data on maintenance needs for industrial machines".

[16] According to the Opinion 05/2014 on Anonymisation Techniques adopted on 10 April 2014 by Article 29 Data Protection Working Party anonymisation "irreversibly" prevents identification of the data subject. Irreversible de-identification is then the goal of any anonymisation process.

[17] *See* Art. 2 and Recital 9 of the Regulation 2018/1807.

According to the purpose limitation principle, which is one of the general principles of the GDPR (Art. 5)[18], personal data must be processed for specified and legitimate purposes and any further processing is lawful only for purposes that are compatible with the original ones. This principle is considered as a "cornerstone of data protection and a prerequisite for most other fundamental requirements" (De Terwangne, 2020, 315). The data processing for a scientific purpose is *a priori* compatible with a primary purpose (Art. 5(b) GDPR, the so-called "presumption of compatibility")[19]. This legislative choice enhances secondary use of personal data for scientific research, but such possibility is not to be taken for granted: a specific, explicit and legitimate secondary purpose is still necessary for the secondary use of personal data to be lawful. In fact, the European Data Protection Supervisor (EDPS) has specified that the presumption of compatibility does not excuse the data controller to conduct a careful analysis of compatibility and implement specific safeguards, such as a data protection impact assessment[20]. Scientific research purposes may correspond to the primary use of the collected data (i.e. data is directly collected for the specific study), or they may correspond to the secondary use (i.e. data is further processed after being initially collected for another purpose), by the controller or, possibly, a different entity. In each case, data processing activities, including storing, accessing and sharing, must be lawful (i.e. lawfulness principle) and comply with the all the other requirements of the GDPR.

Another relevant principle is data minimisation (Art. 5(1)(c) GDPR): personal data should be adequate, relevant and limited to what is necessary in relation to the purpose of the activities. This necessity requirement refers to the quantity and the quality of personal data (De Terwangne, 2020, 317). Data minimisation in a scientific research context is complex to achieve. The same may be true for the application of the storage limitation principle (Art. 5(1)(e) GDPR), which requires that personal data should be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes. Even though the data collected for research purposes may be kept beyond this period, they cannot be stored forever (Ducato, 2020).

[18] The general data protection principles are: lawfulness, fairness, transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity and confidentiality, and accountability.

[19] As an example, a primary purpose may be the healthcare provision.

[20] *See* the "Preliminary Opinion on data protection and scientific research" of the EDPS, adopted on 6 January 2020, pp. 21, 24. Rules on the data protection impact assessment are detailed in Artt. 35-36 of the GDPR.

A scientific project may involve the processing of "data concerning health" which falls amongst special categories of personal data under the GDPR. Data concernig health is "related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status" (Art. 4(15) GDPR). Medical and clinical data, administrative data used for healthcare purposes and genetic data are all included in this broad notion (Bincoletto, 2021, 198-202). The GDPR generally prohibits the processing of particular categories of personal data unless one of the specific exceptions provided in Art. 9(2) GDPR applies. A scientific project processing personal health data uses both common and particular data[21]. Therefore, the researcher should find lawful legal grounds for both categories (Artt. 6 and 9) to comply with the lawfulness principle[22].

Moreover, the GDPR leaves space for Member States to further regulate the processing of data concerning health (Art. 9(4) GDPR). As a result, the researcher should have knowledge of the applicable rules at national level. The Member States' conditions may differ at the detriment of harmonisation at the EU level (Hansen et al., 2021). When a health project is conducted in multiple countries, the European Data Protection Board (EDPB) recommends using the same legal basis, unless national law provides differently[23]. At the same time, this law may establish specific rules that create favourable conditions for scientific research in the medical field.

As regard the exceptions, as well as the legal grounds to process personal data in the medical scientific context, the researcher may choose the data subject's explicit consent pursuant to Art. 9(2)(a) or directly the so-called "research exception" when the processing is necessary for scientific purposes and it is based on Union or Member State law that provides for "suitable and specific measures to safeguard the fundamental rights and the interests of the data subject" (Art. 9(2)(j) GDPR). The consent is the special legal basis for all the clinical trials conducted in the EU according to Article 28 of the Regulation (EU) No 536/2014[24]. In some cases (e.g. post market studies after a

[21] For the purpose of this paper, the term "personal health data" means "data concerning health" in the meaning of the GDPR and "common" data refers to contact details, biographic data, and all the other personal data that are not special.

[22] *See infra.*

[23] *See* the "Document on response to the request from the European Commission for clarification on the consistent application of the GDPR, focusing on health research" of the European Data Protection Board (EDPB), of 2 February 2021, p. 5.

[24] 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 repeal-

clinical trial), where a Union or Member State law provides specific conditions, the research may be carried out under the public health interest exception (Art. 9(2)(i) GDPR) (Comandè, Schneider, 2022a, 577-579)[25].

In the "Preliminary Opinion on data protection and scientific research", adopted on 6 January 2020, the EDPS indicated that Member State law generally requires prior consent from the participant in a research project processing health data and that the "research exception" is operational only when a specific Union or Member States law has been adopted to include conditions and safeguards[26]. In some Member States, law defines the consent as the condition for data processing under the "research exception".

Informed, specific, freely given and unambiguous consent seems hence the main processing ground[27] (Chen et al., 2022). Data subjects also have the right to withdraw their consent anytime. However, forms of broad and dynamic consents are frequently used by the researchers to open the uses of the collected data (Mostert et al., 2018, 52). Such an approach seems partially supported by the interpretation of Recital 33 of the GDPR that states that "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". Such flexibility does not imply that a generic consent for "every research purpose" would be lawful, but it provides the possibility to describe a scientific purpose "in more general terms and for specific stages of a research project that are already known to take place at the outset". Such crucial clarification comes from the "Document on response to the request from the European Commission for clarification on the consistent application of the GDPR, focusing on health research" of the EDPB, of 2 February 2021, which provides some guidance in this context. Relying on the presumption of compatibility does not mean

---

ing Directive 2001/20/EC, O.J. L. 158, 27.5.2014. *See* also the "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))", adopted on 23 January 2019 by the European Data Protection Board.

[25] *See* also the "Preliminary Opinion on data protection and scientific research" of the EDPS, 23: it is difficult "at present, if not impossible, to view a 'substantial public interest' as a basis for processing sensitive data for scientific research purposes" (p. 23). Therefore, the use of the exception of Art. 9(2)(g) GDPR is less probable than the others presented above.

[26] *See* the "Preliminary Opinion on data protection and scientific research" of the EDPS, pp. 14, 17.

[27] It should be specified that the consent under data protection law should be distinguished from the ethical standard for research that requires the collection of the informed consent as a human participant in a research study.

that it is not necessary to comply with Article 9 GDPR or Union or Member State law[28]. When the legal ground chosen is the consent, the will of the data subject seems necessary for the further use of his or her personal data.

The GDPR sets down a special and preferential regime for scientific research in Article 89 (Ducato, 2020) (Schneider, 2020). This provision derogates some data controller's obligations if appropriate safeguards are put in place at Union and Member States levels[29]. A limited derogation also regards the obligation to provide information (Art. 14(5)(b) GDPR). Member State law may even derogate Articles 12-22 GDPR on data subject's rights when the data processing is necessary for general public interests, including public health, and safeguards are put in place (Art. 23)[30].

In conclusion, it may be stressed that the researchers should follow the integrity, confidentiality and accountability principles. The data controller is responsible for and shall able to demonstrate compliance with all the data protection requirements[31]. Allocation of duties and implementation of technical and organisational measures should be defined in advance when planning the research to protect personal data. The obligations of Article 25 of the GDPR on data protection by design and by default play a fundamental role in every data processing (Bincoletto, 2021), including the governance of a scientific project.

At first glance, data protection rules seem to pose a barrier to use and share information. However, as prominent legal scholars already noted, this legal framework "does not hamper but rather encourages data-driven research" (Comandè, Schneider, 2022a, 561). More specifically, a balance with data sharing can be struck. As anticipated, access to data is necessary to innovation and knowledge in the health sector. Pursuing this goal, the following section analyses the current EU framework on open research data and the relevant new legislative initiatives for the health sector.

---

[28] *See* paragraph 22 of the Document.

[29] In particular, Union or Member State law may derogate the following data subject's rights: right to access, to rectification, to restriction, to object.

[30] This possibility has been used during the pandemic to legitimise some exceptional data processing activities. *See* the report of TIPIK Legal of 2021, "Report on the implementation of specific provisions of Regulation (EU) 2016/679" issued for the European Commission, Directorate - General for Justice and Consumers, Unit C.3 Data Protection, and available at https://www.dataguidance.com/sites/default/files/1609930170392.pdf.

[31] Among all the obligations, it may be stressed the importance of the information duties, of enforcing data subject's rights, and the performance of the data protection impact assessment.

## 3. The EU Regime on open research data

Since the early 2000s, the so-called Open Data movement aspires to reduce economic, technological and legal barriers to accessing to information (Guarda, 2021). To the present day, EU rules have been focused on releasing public sector information since public bodies hold an enormous amount of data (Borgesius, Gray, Van Eechoud, 2015). Within this framework, free sharing and free re-use might also involve data collected and used by scientific research.

In 2009, the Open Access Infrastructure for Research in Europe (OpenAIRE) project was launched to create an infrastructure able to share information between research data repositories[32]. The European Open Science Cloud (EOSC) was later promoted by European institutions for the data exchange across borders and scientific disciplines[33]; this initiative benefited researchers, universities and research centers at European level to create "a sort of Internet of data and research services, which must be FAIR, an acronym that stands for Findable, Accessible, Interoperable and Reusable" (Paseri, Varrette and Bouvry, 2021, 128). These principles were firstly elaborated for the re-use of scholarly data (Wilkinson et al., 2016) and then became the guidelines to manage data in scientific research.

The EU policies entail a principle of free movement of research data both for public and private datasets (Comandè, Schneider, 2022a, 567). In the context of publicly-funded research, current rules promote openness both for science and access. The Regulation (EU) 2021/695 establishing Horizon Europe - the Framework Programme for Research and Innovation[34] - defines Open Science as "an approach to the scientific process based on open cooperative work, tools and diffusing knowledge", and Open Access as "online access, provided free of charge to the end user, to research outputs" (Art. 2(5) and (4)). Art. 14 of this Regulation specifies that Open Science implies both the "access to scientific publications" and the "access to research data, including those underlying scientific publications, in accordance with the principle 'as open as possible, as closed as necessary"'. Therefore, the general principle for research data obtained through the Programme is openness,

[32] *See* the official website of the initiative at https://cordis.europa.eu/project/id/643410.

[33] *See* the official websites at https://cordis.europa.eu/project/id/643410 and https://eosc-portal.eu/.

[34] Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe - the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination, and repealing Regulations (EU) No 1290/2013 and (EU) No 1291/2013, O.J. L. 170, 12.5.2021.

but exceptions are possible. These exceptions regard the "legitimate interests" of the researchers that are beneficiaries, meaning "commercial exploitation and any other constraints, such as data protection rules, privacy, confidentiality, trade secrets, Union competitive interests, security rules or intellectual property rights" (Art. 39). Limits to openness are clearly based on existing data protection and intellectual property rules, but also on commercial interests.

The definition of research data was provided for the first time by the Directive (EU) 2019/1024 on Open Data and the re-use of public sector information, or "Open Data Directive"[35], which includes scientific research as a relevant sector. This choice has been considered as a "change of direction" of the EU approach on Open Data (Caso, 2022, 815). Research data are "documents in a digital form, other than scientific publications, which are collected or produced in the course of scientific research activities and are used as evidence in the research process, or are commonly accepted in the research community as necessary to validate research findings and results" (Art. 2).

According to this framework, EU Member States and bodies that fund and conduct public scientific research should adopt national policies for data sharing and re-use of research data. The provisions on research data enhance its role in the data economy and the Digital Single Market (Arisi, 2022). In particular, Art. 10(1) mandates Member States to support the availability of research data by "adopting national policies and relevant actions aiming at making publicly funded research data openly available ('open access policies'), following the principle of 'open by default' and compatible with the FAIR principles". To achieve this goal, Member States should take into account "concerns relating to intellectual property rights, personal data protection and confidentiality, security and legitimate commercial interests" in accordance with the principle of "as open as possible, as closed as necessary".

National policies will be directed to research performing organisations and research funding organisations (Arisi, 2022, 49). On the one hand, Article 10 pushes forward openness "by default" and, when possible, data is made available under open licences (e.g. Creative Commons) and should follow the FAIR principles; on the other hand, pre-existing limits to openness are preserved, including intellectual property rights (of third parties involved in the research), patents, trademarks, registered designs, *sui generis* rights, privacy and data protection, con-

[35] 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, O.J. L. 172/56, 26.6.2019.

fidentiality, national security and legitimate commercial interests (e.g. trade secrets)[36].

Once again data protection law works as limiting open data[37]. This legislative choice considerably restricts the use and re-use of research data. The Open Data Directive explicitly states that it abides to the GDPR (Art. 1(4)), while it also remains limited to the public sector information and it requires national implementation, which may create different open access policies.

New legislative initiatives of the European Commission aim at opening data sharing and access for research purposes. In particular, the recent Regulation (EU) 2022/868 or "Data Governance Act" (DGA)[38] adds conditions for the re-use of certain categories of data held by public sector bodies, including personal data[39]. As regards scientific research, this Act introduces the interesting concept of "data altruism". It is worth quoting in full the notion (Art. 2(16)) that means:

> "the voluntary sharing of data on the basis of the consent of data subjects to process personal data pertaining to them, or permissions of data holders to allow the use of their non-personal data without seeking or receiving a reward that goes beyond compensation related to the costs that they incur where they make their data available for objectives of general interest as provided for in national law, where applicable, such as healthcare, combating climate change, improving mobility, facilitating the development, production and dissemination of official statistics, improving the provision of public services, public policy making or scientific research purposes in the general interest".

Under this definition, research should pursue a general interest. This is the starting point for data sharing, which relies on a very complex infrastructure based on "data altruism organisations" and intermediaries (Comandè, Schneider, 2022b, 776). However, the concept is broad

[36] As argued by Caso, and according to the traditional legal approach, data *per se* can not be the subject of property (i.e. ownership over tangible assets), and of intellectual property. However, this approach appears to be contradicted by legislative and jurisprudential developments which are extending intellectual property rights even to the "basic building blocks of knowledge (data and information)", which should be excluded by the idea-expression dichotomy (Caso, 2022, 818).

[37] *See* also Art. 1(2)(h) that excludes from application other documents. Member States are not mandated to adopt rules related to documents which access is limited at national law level to protect personal data.

[38] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), O.J. L. 152, 3.6.2022.

[39] The categories of data regulated by the Open Data Directive fall outside the scope of the Data Governance Act. *See* Recital 10 of this last Regulation.

and defining the contours of general and private interests in the research context may prove difficult, unless they are specifically grounded and detailed in EU or national law, or they are set by competent public entities who grant access permits. The lack of a clear notion of "general interest" may lead to interpretation issues and thus legal uncertainly (Baloup et al., 2021, 42-43). In addition, the use of the term "general" may lead to confusion in implementation, since the GDPR instead refers to "public interest" (Shabani and Yilmaz, 2022, 131). Questions remain regarding whether privately funded research that pursues a public interest is included. In the context of secondary use of personal health data for scientific projects and public interests, the French approach for the Plateforme des données de santé (Health Data Space) and the Finnish framework for the secondary use of health and social data prove important benchmarks (Bincoletto, Guarda, 2021, 63-64) (Balint and Buki, 2022, 92-96).

The notion of data altruism seems opening the consent of data subject to a "broad consensus for research purposes" (Comandè, Schneider, 2022a, 575). At the same time, the Data Governance Act does not create new legal grounds for data processing and, in the event of conflict, the data protection framework should prevail[40]. The necessity of an explicit consent, therefore, remains. Indeed, Art. 25 of the Data Governance Act provides that the data altruism consent form should ensure that the data subject is able to give consent to and withdraw consent from a specific data processing[41]. In "the EDPB-EDPS Joint Opinion 03/2021 on the Proposal for a regulation of the European Parliament and of the Council on European data governance (Data Governance Act)", of 9 June 2021, the authorities underlined that the GDPR should still apply when the data subject has given consent to the data altruism organisation[42]. The limits of the interpretation of Recital 33 of the GDPR remains applicable[43].

In the interplay between the GDPR and the DGA there may also be a lack of semantic consistency. Although the DGA scope is limited to public sector data, the definitions of Art. 2 of "data holder" (i.e. who has the right to grant access to or to share certain personal data or non-personal data) and "data user" (i.e. who has lawful access to certain personal or non-personal data and has the right to use that data for commercial or non-commercial purposes) should be coordinated with the data processing roles and principles of the GDPR (e.g. of data sub-

[40] *See* Recital 4 of the Data Governance Act.

[41] A model of consent form will be established by the European Commission after consulting the EDPB.

[42] *See* paragraph 161 of the Joint Opinion 03/2021.

[43] Ibidem, paragraph 169.

ject, data controller, data processor and recipient). This inconsistency might lead to difficulties of practical application: the actors on data sharing and their roles under data protection law should always be clear.

With the Proposal for a Regulation on the European Health Data Space (EHDS) the European Commission intends to create a health specific ecosystem[44]. This can be considered the first proposal of a domain-specific common European space (Bincoletto, 2022). Nine common data spaces have been announced by the Commission in the Data Strategy. This new initiative supports both individuals to have more control over their data and the general public while fostering healthcare delivery, research, innovation and policy making. The EHDS Regulation will provide "rules, common standards and practices, infrastructures and a governance framework" on health data for their use and re-use.

The EHDS Proposal has dedicated the entire Chapter IV to the secondary use of electronic health data, including the processing of personal data for research purposes[45]. A "data holder" can make data already collected available for several purposes. In particular, the current version of Art. 34 includes "scientific research related to health or care sectors" amongst the purposes for which electronic health data (i.e. data impacting health, human genetic data, health data registries and clinical trials data) can be processed for secondary use. Some secondary purpose are prohibited[46]. When possible, the access to data will be provided through a governance system ruled by "health data access bodies" at national level. These entities will receive data application requests and, when conditions are met, they will grant data permits. The same bodies will be responsible of informing the data subjects on the legal basis under which access is granted, the measures implemented to protect their rights, the arrangements to exercise data protection rights and the results or outcomes of the projects for which the health data were processed[47]. Data access is also promoted across borders[48]. The legal ground for secondary uses of data may be the compliance with a legal obligation or public interests in combination with the GDPR exceptions and bases (Shabani and Yilmaz, 2022, 133).

[44] Proposal for a Regulation on the European Health Data Space, COM(2022) 197 final, 2022/0140(COD).

[45] According to Art. 2 of the Proposal, "electronic personal data" include "personal electronic health data", which means "data concerning health and genetic data as defined in Regulation (EU) 2016/679, as well as data referring to determinants of health, or data processed in relation to the provision of healthcare services, processed in an electronic form".

[46] *See* Art. 35 of the EHDPS Proposal.

[47] *See* Artt. 36-38 of the Proposal.

[48] *See* the fourth Section of Chapter IV of the Proposal.

Article 40 of the EHDS Proposal is dedicated to "data altruism in health": this requirement explicitly refers to the provision of the Data Governance Act to detail its application in healthcare context. The EHDS Proposal uses the data minimisation and purpose limitation principles to establish that health data should be processed only in line with the data permit obtained from the national bodies[49].

Therefore, it may be argued that the EHDS Proposal aims at opening health data (both personal and non-personal), but it does not deploy the traditional notion of "Open Data". In fact, data may be opened to a specific secondary use on the basis of an authorisation and the scope is not limited to the data held by the public sector. The initiative may boost data access and sharing in the health domain, and it may have a great impact on scientific research. The FAIR principles are quoted in Recital 3 of the act where it is also specified that the new data space should be coordinated with the EOSC.

In the long and detailed "EDPB-EDPS Joint Opinion 03/2022 on the Proposal for a Regulation on the European Health Data Space" of 12 July 2022, the authorities focused on the main issues of the Proposal from a data protection point of view. All the aspects cannot be covered in full detail by this paper. However, as regards the secondary use of health data, it may worth underlining that the EDPB and EDPS noted that this concept is not aligned with the notion of "further use of personal data" of the GDPR. Furthermore, the use of the wording "data holder" is not clear. It should be consistent with the respective concept of the DGA and the definitions and roles of the GDPR. The new regulation should be unambiguous about who has the power and obligation to make health data available for secondary uses, including research. A similar observation is valid with reference to the notion of "data user". The authorities asked for clarification on the interplay between the new possible rules and the specific provisions on the data processing for scientific research established by the GDPR, and the exceptions of Art. 9 particularly (i.e. the legal grounds for data processing)[50].

As pointed out in the draft report of the European Parliament on the Proposal of 10 February 2023, consent is the legal basis for the processing of health data for research uses in many Member States; therefore, processing these data for secondary purposes without the prior consent needs a "significant shift in data protection law as applied" and it creates "an important precedent for further legal acts on

[49] *See* Art. 44.
[50] *See* paragraphs 85-89 of the Joint Opinion 03/2022.

secondary data use"[51]. The Parliament also stated that "a right to a partial or entire opt-out for some or all of the purposes of secondary use should be provided" in order to ensure the right to object (Art. 21(6) the GDPR)[52]. The European Patient Forum has recently proposed a similar "opt-out" mechanism as an amendment to Art. 33 to give patients the choice to control their own health data[53].

Given the detailed set of rules on data protection and open research, it seems vital to discuss the interplay between these two in the following section.

## 4. Tensions between the data protection regime and the open data policies

An unconditional free and open flow of personal data, including data concerning health, even for scientific research purposes, seems not compatible with the principles of the GDPR. In fact, the legislation on Open Data and open research data is without prejudice to data protection requirements. The main challenges that justify the existence of this safeguard may be summarised along the following lines:

– the strict application of the purpose limitation principle, which contrasts with the openness of the purposes in an open data context. This represents the first challenge since purpose limitation is a cornerstone in data protection law. The promotion of re-use conflicts with the purpose limitation rules (Suman, Pierce, 2018, 289). At the same time, the processing for scientific purposes is considered *a priori* compatible with a primary purpose by the GDPR and this mitigates the application of the principle;

– the compliance with the lawfulness principle, which requires the existence of a specific legal ground. The different approach of opening data allows every use when data is opened to the public. The EU

[51] *See* the "Draft Report on the proposal for a regulation of the European Parliament and of the Council on the European Health Data Space" of the Committee on the Environment, Public Health and Food Safety Committee on Civil Liberties, Justice and Home Affairs at the European Parliament, 10.02.2023, 2022/0140(COD).

[52] *See* ibidem pag. 94.

[53] *See* the document of the European Patient Forum of November 2022 "Amendments to the proposal for a European Health Data Space - by the European Patients' Forum (EPF)", available at https://www.eu-patient.eu/contentassets/6d8824eb20224dcb893c12fbc233ec2e/epf—ehds-amendments-final.pdf, pag. 4.

borders and the territorial scope of the GDPR may be overcame by the online upload of the dataset[54];

– the possibility of the revocability of consent always granted by data protection law, while open data appears irrevocable, especially if shared online with many other data. It should however be mentioned that in the research context the withdrawal often does not operate in order to protect the goal of the processing;

– the limitation of the collection and processing grounded on the data minimisation and storage limitation principles, which may be not compatible with the needs of the use and re-use of as many data as possible. Nevertheless, these principles may be mitigated in a research context;

– the compliance with the accountability principle, which requires an effective control and responsibility over data that in the open data scenario disappears or at least is weakened. Releasing opened data should imply that information can be re-used for any purpose and multiple times. Thus, the data controller has less control over the collected data. The data controller would not be able to implement organisational and technical measures if data is disseminated. Data protection by design measures should be implemented in advance, before allowing the re-use. Clear conditions for sharing open research data may alleviate this issue;

– the necessity to inform the data subject on the data processing activities and to adopt the transparency principle. The re-use of data in open format conflicts with this obligation because it is impossible to predict all activities. When personal data is released in an open dataset, the data subject loses control over his or her information (Guarda, 2020, 399);

– the compliance with the specific rules on data concerning health that mandate particular safeguards while processing this category of data because significant economic, psychological and social harms (e.g. employment and social exclusion, stigma, embarrassment and forms of discrimination) may be caused by unauthorised and uncorrected access or sharing of personal health data. Health data may be considered as the most sensitive and intimate personal data. Therefore, personal health data should never be disseminated and only shared under safeguards.

[54] Therefore, even the problem of international trasfer of personal data arises.

As presented in Section 2, data protection rules focus on the notion of legal basis, and in the presence of personal health data a legal basis provides an exception to a general prohibition of processing. While a "pure" open data approach stresses the ability to use and re-use without conditions (i.e. unlimited openness), the GDPR requires an exception to process data for research purposes, which is frequently the data subject's consent. Therefore, personal data may be further processed for a secondary research purpose only on a lawful legal basis. It has been argued that a new consent may be necessary before releasing data as "open research data" (Guarda, 2020, 404).

Thus, it seems possible to achieve two solutions of compromise: anononymisation of personal health data for research purposes to exclude the application of the GDPR, or opening data only for specific research purposes authorised time-to-time under data protection law and in a way that is FAIR.

Firstly, when it proves possible as well as adequate to the scientific research purpose, data could be processed in anonymised form to avoid the application of the GDPR. The data controller anonymises the data and then releases the dataset in an open repository (e.g. Open Science Framework - OSF). Even the DGA encourages the application of anonymisation techniques to share datasets of public sector bodies[55]. However, when choosing anonymisation, there are perplexities related to the techniques applicable, since the data controller assumes the responsibility of their effectiveness in making de-identification irreversible. As already mentioned, it is clear in the literature how difficult it is to create datasets that are effectively anonymised (Podda, Palmirani, 2021). The processing of data rendered anonymised and opened could still present residual risks. In addition, research often cannot resort to anonymised data. In the medical context, the more information the research has on a single subject, the more accurate the treatment responses and patterns analysed will be, and the more useful the analysis will prove for replication, so pseudonymization is more common than anonymisation. For example, only once a predictive model has already been created and tested with personal health data in one project, subsequent research validating it in other scenarios might more easily use non-personal data. It has been pointed out that anonymisation breaks the link between information needed for full traceability of medically relevant research, that it is nearly impossible in biobanking and it is not useful in studies containing stages of follow up of participants (Govarts et al., 2022).

[55] *See* Recitals 9 and 15, and Art. 5(3)(a) on the conditions for re-use.

A second option is creating the conditions to open a specific dataset for a particular secondary scientific research purpose while still complying with the GDPR. There is no "one-size-fits-all" solution that is valid for every dataset: data sharing and access should be managed adopting a data protection by design approach and following the FAIR principles. Data sharing agreements between the former data controller and the research entity may be envisaged, in the context of a research that has a public interest especially. The Proposal of the EHDS may be perceived as the creation of a new framework able to interact "by design" with the data protection rules under an "open research data" paradigm.

The notion of data altruism has great potential in this context, but it should be efficiently coordinated with the GDPR and the provisions already introduced in the DGA. It has been argued that the altruism mechanism requires prior definition of its interplay with the consent of the GDPR and the notion of research purposes in the general interest (Shabani, 2022). More clarity on the meaning of data altruism and then on this "opt-in" option in the research sector is therefore required. It may also be suggested that the altruistic purposes detailed in the consent form may be limited to specific areas of medical or biomedical research that pursue public interests, or that consent may be given to a series of scientific projects that have a similar general purpose. Creating technical intermediaries able to contact the data subject to ask for an explicit consent is another possible solution, which needs a robust oversight by designated public entities. Whether the legal ground for the data sharing is a legal obligation or a general/public interest according to a provision of the EHDS Regulation or to a national law, further safeguards and organisational and technical measures should protect personal health data anyway. Member States law may create conditions applicable to the scientific research that uses personal health data, as already provided in the past (Slokenberga, 2022). The conditions for data sharing for scientific research projects "in the general interest" should be precise and all the possible inconsistencies between "data frameworks" should be avoided.

## 5. Conclusive remarks

This paper shows that tensions exist between the data protection regime and the policies of opening data. The use of personal health data for scientific research purposes shall comply with the data protection framework to safeguard fundamental rights and freedoms of the data subjects. At the same time, opening this data is indispensable for sci-

entific progress and can benefit the entire community in terms of an improved public health.

Personal data protection requires to comply with all the GDPR principles, including purpose limitation (pre-defining the purpose of the processing), lawfulness (having a lawful legal ground for every activity) and accountability (overall protection and implementation of appropriate and effective safeguards, possibly by design). Understanding whether it is possible to open up the data covers both the case where data are collected to conduct research (primary purpose) and the researchers want to make the same data available to others, as well as the case where a researcher need to access data collected for other purposes in the past (secondary purpose).

Emerging new strategies and rules suggest that openness will increase in the near future. The notions of research data of the Open Data Directive and of data altruism of the DGA may play a pivotal role. One of the main challenges will be their alignment with the relevant horizontal and sectoral EU and national legislation.

As regards the EHDS Proposal, it may be suggested that the definitions and concepts should be very aligned with the GDPR. Despite the potential to boost data sharing, the act needs several amendments to be compliant with the data protection rules. Confusions needs to be avoided as regarding data access and re-use.

The possibility that Member States further regulate data concerning health and processing for scientific research purposes may also be obstacle for harmonisation and for the development of the EHDS (Kiseleva, de Hert, 2022, 26).

Regulating accessibility and openness requires a reasonable approach that takes into account the existing legal requirements. Besides anonymisation, the application of a data protection by design approach on a case-by-case basis (to be preferred to a one-size-fits-all solution) allows data management practices that open the collected personal health data for specific scientific projects.

Finally, in the near future the EDPB will provide important guidance in this context. In fact, the "Guidelines on processing of data for medical and scientific research purposes" and guidelines on anonymisation have been included in the EDPB Work Programme 2023-2024 adopted on 22 February 2023[56]. Therefore, we should be mindful of what is to come.

[56] *See* the official website of the authority at https://edpb.europa.eu/.

## Acknowledgements

The author would like to thank the two anonymous reviewers for the insightful comments on an earlier version of this paper.

## References

Arisi, M. (2022), *Open Knowledge. Access and Re-use of Research Data in the European Union Open Data Directive and the Implementation in Italy*, Italian Law Journal, No. 1, pp. 33-73.

Balint, F. and Buki, B. (2022), *Three Ways of Secure Data Reusability in Europe: German Research Data Centres, Finnish Findata and the French Secure Access Data Centre*, ELTE Law Journal, No. 1, pp. 81-108.

Baloup J. et al (2021), *White Paper on the Data Governance Act*, available at: SSRN: https://ssrn.com/abstract=3872703 (accessed 12 May 2023).

Besançon, L. et al. (2021), *Open science saves lives: lessons from the COVID-19 pandemic*, BMC Medical Research Methodology, No. 1, pp. 1-18.

Bincoletto, G. (2022), *The EDPB-EDPS Joint Opinion on the Commission Proposal for a Regulation on the European Health Data Space: Key Issues to be Considered in the Legislative Process*, European Data Protection Law Review, No. 3, pp. 398-404.

Bincoletto, G. (2021), *Data Protection by Design in the E-Health Care Sector*, Nomos Verlagsgesellschaft mbH & Co. KG, Baden-Baden.

Bincoletto, G. and Guarda, P. (2021), *A proactive GDPR-compliant solution for fostering medical scientific research as a secondary use of personal health data*, Opinio Juris in Comparatione, No. 1, pp. 43-76.

Caso, R. (2022), *Open Data, ricerca scientifica e privatizzazione della conoscenza*, Diritto dell'informazione e dell'informatica, No. 4/5, pp. 815-836.

Caso, R. (2020), *La scienza non sarà più la stessa. Più condivisione, cooperazione e solidarietà dopo il Covid-19?*, BioLaw Journal, No. 1s, pp. 617-622.

Caso, R., Binda, F. (2020), *Il diritto umano alla scienza aperta*, In G. Perrone, M. Perduca (Eds.), "Così san tutt3 - Diritto alla Scienza, istruzioni per l'uso", Fandango Libri, Roma, pp. 44-52.

Chen, J., S. Dove, E., Bhakuni, H. (2022), *Explicit consent and alternative data protection processing grounds for health research*, In G. Perrone, M. Perduca (Eds.), "Research Handbook on EU Data Protection Law", Edward Elgar Publishing, Cheltenham, pp. 474-502.

Comandè, G. (2022) (Eds.), *Elgar Encyclopedia of Law and Data Science*, Edward Elgar Publishing, Cheltenham.

Comandè, G. and Schneider, G. (2022a), *Differential Data Protection Regimes in Data-Driven Research: Why the GDPR is More Research-Friendly Than You Think*, German Law Journal, Vol. 23 No. 4, pp. 559-596.

Comandè, G. and Schneider, G. (2022b), *It's time: Leveraging the GDPR to shift the balance towards research-friendly EU data spaces*, Common Market Law Review, Vol. 59 No. 3, pp. 739-776.

Ducato, R. (2020), *Data protection, scientific research, and the role of information*, Computer Law & Security Review, Vol. 37, p. 105412.

Guarda, P. (2021), *Il regime giuridico dei dati della ricerca scientifica*, Editoriale Scientifica, Napoli.

Guarda, P. (2020), *Free data?: open science in the age of personal data protection*, In H.J. Rooksby (Ed.), "Research Handbook on Intellectual Property and Technology Transfer", Edward Elgar Publishing, Cheltenham, pp. 391-410.

Hallinan, D. (2021), *Protecting Genetic Privacy in Biobanking through Data Protection Law*, Oxford University Press, Oxford.

Hansen, J. et al. (2021), *Assessment of the EU Member States' rules on health data in the light of GDPR*, Publication Office of the European Union, Luxembourg.

Hansson, M. G. (2021), *Striking a Balance Between Personalised Genetics and Privacy Protection from the Perspective of GDPR*, In S. Slokenberga et al. (Eds.), "GDPR and Biobanking", Springer, Cham, pp. 31-42.

De Terwangne, C. (2020), *Chapter II Principles (Articles 5–11). Article 5. Principles relating to processing of personal data*, In Kuner C. et al., (Eds.), "The EU General Data Protection Regulation (GDPR): A Commentary", Oxford University Press, Oxford, pp. 309-397.

Borgesius, F.Z., Gray, J. & Van Eechoud, M. (2015), *Open Data, Privacy, and Fair Information Principles: Towards a Balancing Framework*, Berkeley Technology Law Journal, Vol. 30 No. 3, pp. 2073-2132.

Kiseleva, A., and de Hert, P. (2021), *Creating european health data space: obstacles in four key legal area*, European Pharmaceutical Law Review, Vol. 5 No. 1, pp. 21-36.

Kuner, C., Bygrave, L.A., Docksey, C. (Eds.) (2020), *The EU General Data Protection Regulation (GDPR): a commentary*, Oxford University Press, Oxford.

Laurie, G., Dove, E., Ganguli-Mitra, A., McMillan, C., Postan, E., Sethi, N., & Sorbie, A. (Eds.) (2021), *The Cambridge Handbook of Health Research Regulation*, Cambridge University Press, Cambridge.

Madir, J. (Ed.) (2020), *Healthtech. Law and Regulation*, Edward Elgar Publishing, Cheltenham.

Govarts, E. et al. (2022), *Position paper on management of personal data in environment and health research in Europe*, Environment International, Vol. 165, p. 107334.

Margoni, T., Caso, R., Ducato, R., Guarda, P. (2017), *Open Access, Open Science, Open Society*, In F. Loizides and B. Schmidt (Eds.), "Positioning and Power in Academic Publishing: Players, Agents and Agendas", Springer, Cham, pp. 75-86.

Mostert, M. et al. (2018), *From Privacy to Data Protection in the EU: Implications for Big Data Health Research*, European Journal of Health Law, Vol. 25, pp. 43-55.

Narayanan, A. and Shmatikov, V. (2008), *Robust De-anonymization of Large Sparse Datasets*, In The Institute of Electrical and Electronics Engineers, Inc. (Ed.), "2008 IEEE Symposium on Security and Privacy", IEEE Computer Society, Los Alamitos, pp. 111-125.

Ohm, P. (2010), *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, UCLA Law Review, Vol. 57, pp. 1701-1777.

Paseri, L., Varrette, S., Bouvry, P. (2021), *Protection of Personal Data in High Performance Computing Platform for Scientific Research Purposes*, In N. Gruschka, L.F.C. Antunes, K. Rannenberg, P. Drogkaris (Eds.), "Privacy Technologies and Policy. APF 2021", Springer, Cham, pp. 123-142.

Podda, E., Palmirani, M. (2021), *Inferring the Meaning of Non-personal, Anonymized, and Anonymous Data*, In V. Rodríguez-Doncel, M. Palmirani, M. Araszkiewicz, P. Casanovas, U. Pagallo, G. Sartor, G. (Eds.), "AI Approaches to the Complexity of Legal Systems XI-XII", Springer, Cham, pp. 269-282.

Purtova, N. (2018), *The law of everything. Broad concept of personal data and future of EU data protection law*, Law, Innovation and Technology, Vol. 10, No. 1, pp. 40-81.

Schneider, G. (2020), *Health Data Pools under European Policy and Data Protection Law: Research as a New Efficiency Defence?*, Journal of Intellectual Property, Information Technology and Electronic Commerce Law, Vol. 11 No. 1, pp. 49-67.

Shabani, M. (2022), *Will the European Health Data Space change data sharing rules?*, Science, Vol. 375 No. 6587, pp.1357-1359.

Shabani, M. and Yilmaz, S. (2022), *Lawfulness in secondary use of health data. Interplay between three regulatory frameworks of GDPR, DGA & EHDS*, Technology and Regulation, Vol. 2022, pp. 128-134.

Shabani, M., Thorogood, A., & Murtagh, M. (2021), *Data Access Governance*, In G. Laurie, E. Dove, A. Ganguli-Mitra, C. McMillan, E. Postan, N. Sethi, et al. (Eds.), "The Cambridge Handbook of Health Research Regulation", Cambridge University Press, Cambridge, pp. 187-196.

Slokenberga, S. (2022), *Scientific research regime 2.0? Transformations of the research regime and the protection of the data subject that the proposed EHDS regulation promises to bring along*, Technology and Regulation, Vol. 2022, pp. 135-147.

Slokenberga, S., Tzortzatou, O., & Reichel, J. (Eds.) (2021), *GDPR and Biobanking: Individual Rights, Public Interest and Research Regulation across Europe*, Springer International Publishing, Cham.

Stalla-Bourdillon, S., & Knight, A. (2017), *Anonymous data v. personal data - a false debate: an EU perspective on anonymization, pseudonymization and personal data*, Wisconsin International Law Journal, Vol. 34 No. 2, pp. 284-322.

Suman, A. & Pierce, R. (2018), *Challenges for Citizen Science and the EU Open Science Agenda under the GDPR*, European Data Protection Law Review, Vol. 4 No. 3, pp. 284-295.

Véliz, C. (2019), *Medical Privacy and Big Data: : A Further Reason in Favour of Public Universal Health-Care Coverage*, In "Philosophical Foundations of Medical Law", Oxford University Press, Oxford, pp. 306–C21.P81.

Wilkinson, M.D. et al (2016), *The FAIR Guiding Principles for scientific data management and stewardship*, available at: https://pubmed.ncbi.nlm.nih.gov/26978244/ (accessed 12 May 2023).