

jipitec

1 | 2020

Volume 11 (2020)
Issue 1 ISSN 2190-3387

Editorial

by Chris Reed

Articles

Cutting Back Patent Over-Enforcement – How to Address Abusive Practices Within the EU Enforcement Framework
by Alain Strowel and Amandine Léonard

Unscrewing the Future: The Right to Repair and the Circumvention of Software TPMs in the EU
by Anthony D. Rosborough

Health Data Pools under European Policy and Data Protection Law: Research as a New Efficiency Defence?
by Giulia Schneider

Free Speech by Design for the Copyright DSM Directive
by Maxime Lambrecht

Book Reviews

Jacques de Werra (ed.), *Accords de technologie/ Technology Transactions*
by Pedro Roffe

Journal of
Intellectual Property,
Information Technology,
and Electronic Commerce
Law

Editors:

Thomas Dreier
Axel Metzger
Gerald Spindler
Lucie Guibault
Miquel Peguera
Séverine Dusollier
Chris Reed
Karin Sein

www.jipitec.eu

Jipitec

Journal of Intellectual Property,
Information Technology and
Electronic Commerce Law

Volume 11 Issue 1 April 2020

www.jipitec.eu

contact@jipitec.eu

A joint publication of:

Prof. Dr. Thomas Dreier, M. C. J. (NYU)

KIT - Karlsruher Institut für Technologie,
Zentrum für Angewandte
Rechtswissenschaft (ZAR),
Vincenz-Prießnitz-Str. 3,
76131 Karlsruhe Germany

Prof. Dr. Axel Metzger, LL. M. (Harvard)

Humboldt-Universität zu
Berlin, Unter den Linden 6,
10099 Berlin

Prof. Dr. Gerald Spindler

Dipl.-Ökonom, Georg-August-
Universität Göttingen,
Platz der Göttinger Sieben 6,
37073 Göttingen

Karlsruhe Institute of Technology,
Humboldt-Universität zu Berlin
and Georg-August-Universität
Göttingen are corporations under
public law, and represented by
their respective presidents.

Editors:

Thomas Dreier

Axel Metzger

Gerald Spindler

Lucie Guibault

Miquel Peguera

Séverine Dusollier

Chris Reed

Karin Sein

Board of Correspondents:

Graeme Dinwoodie

Christophe Geiger

Ejan Mackaay

Rita Matulionyte

Giovanni M. Riccio

Cyrill P. Rigamonti

Olav Torvund

Mikko Välimäki

Rolf H. Weber

Andreas Wiebe

Raquel Xalabarder

Editor-in-charge for this issue:

Chris Reed

Technical Editor:

Lydia Förster

ISSN 2190-3387

Funded by



Deutsche Gesellschaft für
Recht und Informatik e.V.

Table Of Contents

Editorial

by Chris Reed 1

Articles

Cutting Back Patent Over-Enforcement – How to Address Abusive
Practices Within the EU Enforcement Framework
by Alain Strowel and Amandine Léonard 3

Unscrewing the Future: The Right to Repair and the
Circumvention of Software TPMs in the EU
by Anthony D. Rosborough 26

Health Data Pools under European Policy and Data Protection
Law: Research as a New Efficiency Defence?
by Giulia Schneider 49

Free Speech by Design for the Copyright DSM Directive
by Maxime Lambrecht 68

Book Reviews

Jacques de Werra (ed.), Accords de technologie/ Technology
Transactions
by Pedro Roffe 95

Health Data Pools under European Policy and Data Protection Law: Research as a New Efficiency Defence?

by **Giulia Schneider***

Abstract: The increasing employment of artificial intelligence and machine learning in the biomedical sector as well as the growing number of partnerships aimed at pooling together different types of digital health data, stress the importance of an effective regulation and governance of data sharing in the health and life sciences. This paper explores the emerging economic reality of health data pools from the perspective of European Union policy and law.

The goal of the study is to validate the role of the internal market integration objective in the data protection framework of special categories of data, and thus to unveil the alignment of the General Data Protection Regulation's research exemption with the broader policy goals of the Digital Single Market Strategy.

After having described the phenomenon of health data pools as a primary means to conduct research in digital health markets, the study first contextualizes health data sharing practices at European policy level, with specific reference to the Digital Single Market Strategy. Here, both the digital health sector and the free-flow of information are emerging as strategic areas of European intervention.

Against this backdrop, the second section will enquire the regulatory framework regarding the processing of special categories of data for research purposes under the General Data Protection Regulation. As will be demonstrated, this framework partly disavows fundamental rights protection objectives, in order to promote research based on health data and related market objectives.

Keywords: Data sharing; Research; Innovation; Data protection; Digital health

© 2020 Giulia Schneider

Everybody may disseminate this article by electronic means and make it available for download under the terms and conditions of the Digital Peer Publishing Licence (DPPL). A copy of the license text may be obtained at <http://nbn-resolving.de/urn:nbn:de:0009-dppl-v3-en8>.

Recommended citation: Giulia Schneider, Health Data Pools under European Policy and Data Protection Law: Research as a New Efficiency Defence?, 11 (2020) JIPITEC 49 para 1.

A. Introduction and Outline of the Study

1 The increasing employment of artificial intelligence and machine learning in the biomedical sector as well as the growing number of partnerships aimed at pooling together different types of digital health data, stress the importance of an effective regulation and governance of data sharing in the health and life sciences. This paper explores the emerging economic reality of health data pools from the perspective of European Union policy and law. The goal of the study is to validate the role of the internal market integration objective in the data protection framework of special categories of data, and thus to unveil the alignment of the General Data Protection

Regulation's research exemption as a ground for the processing of special categories of data with the broader policy goals of the Digital Single Market Strategy.

2 Innovation in health-related markets, such as the ones of medical devices and pharmaceuticals is growingly occurring through the door of digitisation and datification courses¹. This means that in the

* Research Fellow at Lider-Lab, Sant'Anna School of Advanced Studies, Pisa.

1 This is well expressed by William Nicholson Price II, 'Black Box Medicine' (2015) 28, 2 Harvard Journal of Law & Technology, 420, 422, affirming that "black-box medicine

algorithm-driven economy highly complex data-sets as well as highly sophisticated analytical techniques are needed in order to achieve innovation in health-related markets².

- 3 Traditional actors in the healthcare setting, such as pharmaceutical companies or public healthcare providers, lack of the needed information-technological expertise. They are thus increasingly looking for the support of big data companies, which own mass amounts of users' data, who have the standard technical infrastructure in order to run more sophisticated experiments and thus provide prompter clinical responses. On the other hand, big data companies entering health markets need the more sophisticated health-related data and the expertise traditional stakeholders in the healthcare sector have.
- 4 As a result of the matching between these different economic interests, the conduction of healthcare research is starting to evolve around a complex architecture, where courses of biomedical innovation are driven by new forms of collaborative networks³ between high-tech companies, and traditional stakeholders in the health sector such as pharmaceutical companies and public health providers⁴. These collaborations' primary goal relates to the sharing of different types of health data.

relies principally on pure information goods: collected data, patterns discovered within that data, and validation of those patterns”.

- 2 The fact that the processing and exploitation of complex datasets is key for the success and commercial value of companies acting in digital markets is stressed by Karl-Heinz Fezer, 'Data Property of the People-An Intrinsic Intellectual Property Law Sui Generis Regarding People's Behavior-generated Informational Data' (2017) *Zeitschrift für Geistiges Eigentum*, 356, 356-357, stating that “in the reality of the market, behaviour-generated informational data represents a tradable commodity and crucial asset in a booming industry in the digitized world”.
- 3 The expression is taken from Luis M. Camarinha-Matos and Hamideh Afsarmanesh, 'Collaborative Networks-Value Creation in a Knowledge Society' in: Kesheng Wang and George L. Kovacs and Michael Wozny and Minglun Fang (eds.), *Knowledge Enterprise: Intelligent Strategies in Product Design, Manufacturing, and Management* (Springer, 2006) 26-40.
- 4 From a more general perspective, not strictly related to the medical sector, the emergence of new collaboration scenarios characterising high technology markets, is well highlighted by Giuseppe Colangelo, *Mercato e cooperazione tecnologica. I contratti di patent pooling* (Giuffrè- Quaderni di Aida, 2008) 32 ff.

These sharing practices are giving rise to outright “health data ecosystems”⁵.

- 5 Digital health data represent a highly scientifically valuable asset, the accessibility and the processing of which is ever more becoming essential for research and market innovation purposes in the field of digital health. Economic advancements in this sector are in turn believed to promisingly heighten the standard of health overall enjoyed.
- 6 Health data availability is indeed believed to improve and fasten the design of digital health products, in terms of optimisation and personalisation of the manufacturing processes and with related gains in terms of quality of the resulting products⁶.
- 7 In these regards, according to a growing strand of the literature, regulatory incentives and a correspondent legislative action are needed in order to advance research and innovation in the field of health through the aggregation of differently owned datasets⁷.

5 In this regard, some strand of the literature has referred to “health data ecosystems” in order to describe the “technical and social arrangements underpinning the environments in which health data is generated, analysed, shared and used”. Sonja Marjanovic and Ioana Ghiga-Miaoqing Yang and Anna Knack, 'Understanding Value in Health Data Ecosystems-A Review of Current Evidence and Ways Forward' (Rand, 2017) 1 online available at <https://www.rand.org/pubs/research_reports/RR1972.html>. Emphasis added. Similarly, also Effy Vayena and Alessandro Blasimme, 'Biomedical Big Data: New Models of Control over Access, Use and Governance' (2017) 14 *Bioethical Enquiry*, 501, 503, where the Authors highlight “the interdependence of the actors and processes that rely on the production and circulation of data as a key resource for their respective activities”.

6 Björn Lindqvist 'Competition and Data Pools' (2018) *Journal of European Consumer and Market Law*, 146, 147-148.

7 Arti K. Rai, 'Risk Regulation and Innovation: the Case of Rights-Encumbered Biomedical Data Silos' (2017) 92, 4 *Notre Dame Law Review*, 101 ff.; Rebecca S. Eisenberg and Arti K. Rai, 'Harnessing and Sharing the Benefits of State-Sponsored Research: Intellectual Property Rights and Data Sharing in California Stem's Cell Initiative' (2006) 21 *Berkeley Technology Law Journal*, 1187, 1196-1199. Against this backdrop, the proposed legal incentives are both of private nature, as the establishment of a right to property over health data and the creation of public funders resource creation exercising informal or formal regulatory power to promote data pooling. See, e.g., Jorge L. Contreras, 'Leviathan in the Commons: Biomedical Data and the State' in: Katherine J. Strandburg- Michael J. Madison- Brett M. Frischmann (ed.), *Governing Medical Knowledge Commons* (Cambridge University Press, 2017) 9-18.

- 8 The particularly sensitive nature of the data being shared in the course of digital health research projects renders innovation driven by health data a highly challenging regulatory matter. Innovation and broader public health gains, respectively linked to businesses' fundamental rights to conduct business and to patients' fundamental right to health, need to be carefully outweighed against data protection and discrimination concerns, equally protected as fundamental rights within the European Union.
- 9 Under these premises, the first part of the study theoretically assesses pooling practices as a means of concentrating high-technology resources and stirring innovation in the life sciences sector. In these regards, data pools are considered an evolution of patent pools in the digital economy.
- 10 At a European policy level, health data pools for research purposes are strongly promoted within the Digital Single Market Strategy, being related to both the digital health sector and the free-flow of information initiative.
- 11 Against this backdrop, the second section will enquire the regulatory framework regarding the processing of special categories of data for research purposes under the General Data Protection Regulation. A careful examination of the research exemption under arts. 9(2) lett. j); 5(1) lett. b); 6(4) and 89 GDPR applicable to health data as special categories of data reveals that data-driven health research activities are enabled and promoted under the reformed European data protection law.
- 13 As a result of the digitisation and datification of health research assets, it seems that the "thicket" problem has come to extend well beyond the patent protection of final products and is increasingly affecting the research valuable information that stands behind final products. Such information has become an increasingly strategic asset in the dynamics of competition in the pharmaceutical sector and has been thus progressively encumbered with property-based rights¹⁰. This has triggered the need to expand the range of protection tools employed by originators involved in health research endeavours.
- 14 In addition to trade secret protection and regulatory exclusivities traditionally guarding clinical trials data, also copyright and database protection are emerging as important instruments for shielding collected and processed health data, as well as automatically-generated health inferences and predictions¹¹. Along these lines, also the technological infrastructure employed for the processing and the generation of such data finds legal protection under both patent and copyright regimes over software¹².
- 15 These kinds of information-based protections insisting over digital health data all share the underlying function of protecting digital businesses' competitive advantage deriving from their investments in the collection and production of information. Through the above-mentioned intellectual property tools, and through their *direct* or *indirect* secrecy outcomes, companies' valuable R&D information is gradually shielded from the

B. The Problem of Data Thickets in Digital Health Research

- 12 Traditionally, in the pharmaceutical sector, "patent thickets"⁸, consisting in a bundle of different and intersecting property rights over technology assets, have been regarded as one of the main causes of the freezing of socially-valuable down-stream innovation⁹.

8 For a general assessment of the issue, Carl Shapiro, 'Navigating the Patent Thicket: Cross-Licences, Patent Pools and Standard Setting' in: Adam B. Jaffe, Josh Lerner and Scott Stern, *Innovation Policy and the Economy*, vol. 1 (MIT Press 2001)119 ff. See also Jonathan Barnett, 'From Patent Thickets to Patent Networks: the Legal Infrastructure of the Digital Economy' (2014) *Jurimetrics*, 55 ff., arguing that patent pools and other cross-licensing structures overcome problems of patent thickets and related inefficiencies.

9 Michael A. Heller and Rebecca S. Eisenberg, 'Can Patents Deter Innovation? The Anticommons in Biomedical Research' (1998) 20 *Science* 698; Arti K. Rai, 'Fostering

Cumulative Innovation in the Biopharmaceutical Industry: The Role of Patents and Antitrust' (2001) 16 *Berkeley Technology Law Journal* 813. Talking about 'blocking patents' also Robert Merges, 'Intellectual Property Rights and Bargaining Breakdown: The Case of Blocking Patents' (1994) 62 *Tennessee Law Review*, 75, 81-82. See also Arti K. Rai, 'Regulating Scientific Research: Intellectual Property Rights and the Norms of Science' (1999) 94 *Northwestern Law Review*, 77 ff.

10 Johanna von Braun and Meir P. Pugatch, 'The Changing face of the Pharmaceutical Industry and Intellectual Property Rights' (2005) *The Journal of World Intellectual Property*, 599 ff.

11 Highlighting this point, Giulia Schneider and Giovanni Comandè, 'Regulatory Challenges of Data Mining Practices: The Case of the Never-ending lifecycles of 'Health Data'' (2018) 25 *European Journal of Health Law*, 2018, 284 ff.

12 On the issue see Scott Hensley, 'Software Will Play Key Role in Future Genome Research', (14 February 2001) *Wall Street Journal* <<https://www.wsj.com/articles/SB982100274706275947>>.

free-riding threats of the public domain¹³. Relying on these tools, digital companies can control and limit access over health information- as it happens with the database right and the copyright- or secretize this same information- as it is the case of trade secrets. These different forms of protection over scientific digitised data frequently overlap and create a layered regime of protection over the results of research endeavours, variedly securing scientifically precious information¹⁴. In this perspective, both overlapping and adjacent rights over biomedical data leads to a situation of strict control by the initial rights' holders over upstream technology, *i.e.* scientific data and the technical processing infrastructure¹⁵.

- 16 In addition to legal measures, also factual and technical measure can further enclose companies' research data silos¹⁶. Technical measures of protection have both the effect of factually stretching the limitations on the scope of exclusivities set by the law¹⁷ and, even more interestingly, of factually controlling resources that would not be eligible of protection from both the perspective of objective requirements- as the originality requirement under copyright or the substantial investment requirement under database protection-, and subjective requirement, because the subjects who enact these measures is not the originator of the resource. This means that a specific resource can be appropriated by a player through technical protection measures even if the resource has been originally generated by another company¹⁸.

The above-traced scenario reflects the emergence of a data "thicket" problem, freezing competitors' capacities to compete at a phase that goes well before the marketization of the final product and relates to the previous stage of research over the product itself¹⁹.

- 17 Companies' data "silos" have been strongly criticised in the literature, observing how the excessive control over scientific information gives rise to a situation of "innovation bundling" for which "neither the invention nor the complements can be reasonably developed" without access to the protected information²⁰. This appears to hold true especially in the digital health sector²¹, where the aggregation of different datasets and the statistical insights that result from the combined datasets are becoming a precondition for a faster development and thus a faster marketization of new health-related products and services. For these purposes, the needed correlations and predictions are the more accurate and precise, the bigger the aggregated datasets are.
- 18 Hence, the research and innovation driven by the aggregation of different types of data risks to be obstructed by the existence of different rights over different types of datasets: pharmaceutical companies, for example, have control of traditional clinical trials data whereas digital companies cover with trade secrets scientifically valuable "runaway data".
- 19 The fragmentation of scientific knowledge together with the resulting erosion of publicly available research resources, thus risks to transform the relationship between intellectual property
- 20 This is point is widely raised in the literature, Rai (n [9]) 813; Jerome H. Reichman and Paul F. Uhlir, 'A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment' (2003) *Law & Contemporary Problems*, 315, 402-408. See also Nicholson Price II (n [1]) 447-448, underlining how "keeping data secret" in the area of health research "may significantly hamper the development of black-box medicine. Secrecy slows cumulative innovation and promotes duplicative investment".
- 21 Similarly, Arti K. Rai, 'The Information Revolution Reaches Pharmaceuticals: Balancing Innovation Incentives, Cost, and Access in the Post-Genomic Era' (2001) *University of Illinois Law Review*, 173 ff.

13 James Boyle, *The Second Enclosure Movement and the Construction of the Public Domain* (2003) 66 *Law and Contemporary Problems* 33 ff., and more generally see Joseph E. Stiglitz, *Knowledge as a Public Good in: Inge Kaul and Isabelle Grunberg and Marc Stern, Global Public Goods: International Cooperation in the 21st Century* (Oxford Scholarship Online, 2003) 75 ff.

14 Rai, (n 7) 106-112.

15 *Ibid.*, 102.

16 Reto M. Hilty, 'Intellectual Property and Private Ordering' in: Rochelle Dreyfuss and Justine Pila, *The Oxford Handbook of Intellectual Property Law* (Oxford University Press, 2018) 898 ff.

17 *Ibid.*, 891.

18 Stressing this point Nadia Purtova, 'Health Data for Common Good: Defining the Boundaries and Social Dilemmas of Data Commons', in Ronald Leenes, Nadezhda Purtova and Samantha Adams, *Under Observation: The Interplay Between eHealth and Surveillance* (Springer, 2017), 177, 205.

19 William Nicholson Price II, *Expired Patents, Trade Secrets and Stymied Competition* (2017) 92, 4 *Notre Dame Law Review*, 1611, 1613.

20 This is point is widely raised in the literature, Rai (n [9]) 813; Jerome H. Reichman and Paul F. Uhlir, 'A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment' (2003) *Law & Contemporary Problems*, 315, 402-408. See also Nicholson Price II (n [1]) 447-448, underlining how "keeping data secret" in the area of health research "may significantly hamper the development of black-box medicine. Secrecy slows cumulative innovation and promotes duplicative investment".

21 Similarly, Arti K. Rai, 'The Information Revolution Reaches Pharmaceuticals: Balancing Innovation Incentives, Cost, and Access in the Post-Genomic Era' (2001) *University of Illinois Law Review*, 173 ff.

protection and innovation in digital health research from a “direct” to an “inverse” proportionality relationship²².

- 20 The outcome of this changed scenario is the emerging need of firms to mediate between the possibility to successfully claim exclusivity rights over technological information and the preservation of innovation courses’ fruitfulness²³.

C. The Phenomenon of Data Pools in Digital Health Research

- 21 Concrete organisational responses to the rights’ and resources’ dispersion affecting scientific health information are to be found in collaboration schemes based on data sharing between different actors in the field of medical research. Information alliances achieved through the pooling of intellectual property rights and the establishment of coordination architectures over research patterns are capable -if well designed- to overcome scientific information silos hurdles in a pro-competitive manner²⁴ and thus advance innovation in digital health markets. Under these premises, aggregation of data in pools is to be seen as a direct reaction to the problem of data “thickets” and the precondition of technological innovation in the digital health sector²⁵.
- 22 Pooling practices as a means of concentrating high-technology resources and stirring innovation

in health-related markets, is a traditionally well-known phenomenon. Patent pooling schemes have been indeed largely used in the pharmaceutical sector²⁶.

- 23 They enable the licensing of complementary patents by means of a single agreement and at a standard royalty fee, with the related benefits in terms of cost cuts²⁷. At its very essence, patent pools are a form of technological cooperation between different right owners willing to speed up the process of cumulative innovation²⁸. Assembling together technology assets enables companies to put themselves together to remain at the forefront of information technology developments²⁹, through incentivising coordination mechanisms among participants and the prevention of opportunistic free-riding conducts³⁰.
- 24 Similarly to patent pools, also research cooperation initiatives based on the sharing of health data, imply the licensing of different datasets to a central administrator, who exploits the full potential of the aggregated data through data analytics technologies³¹. As with patent pools these kind of agreements reduce transaction costs related to data collection and processing operations³² and enable to aggregate a large quantity of data, generating more precise and accurate correlations and predictions.

22 This is confirmed by some economics studies, which have framed the relationship between intellectual property law and innovation as an “inverted-U relationship”. So Yuichi Furukawa, *Intellectual Property Protection and Innovation: an Inverted-U Relationship* (2010) *Economics Letters*, 99-101.

23 Colangelo (n 4) 4.

24 Stressing this point with regards to research on genetic data, Turna Ray, ‘Genomic Data Sharing Variant Gains Support. Collaboration Seen as a Key to Interpretation Challenge’ (2 May 2016) *Genome Web*, 2 <<https://www.genomeweb.com/informatics/genomic-variant-data-sharing-gains-support-collaboration-seen-keyinterpretation#.XMrTU5MzYb0>>. Rai (n [9]), 845.

25 Michael Mattioli, *The Data Pooling Problem* (2017) 32 *Berkeley Technology Law Journal* 179, 187, stating that in an information-based economy, incentivising the combination of different large datasets owned by different companies or institutions could serve similar innovation goals to the ones promoted by the patent system in a product-based economy.

26 For an overall assessment see Jorge A. Goldstein, ‘Critical Analysis of Patent Pools’ in: Geertrui Van Oweralle, *Gene Patents and Collaborative Licensing Models: Patent Pools, Clearinghouses, Open Source Models and Liability Regimes* (Cambridge University Press, 2009) 50. See also Robert P. Merges, ‘Institutions for Intellectual Property Transactions: the Case of Patent Pools’ in: Rochelle C. Dreyfuss-Diane Leenheer Zimmermann-Harry First (ed.), *Expanding the Boundaries of Intellectual Property: Innovation Policy for the Knowledge Society* (Oxford University Press, 2001) 123 ff.

27 For an empirical demonstration of the reduction of transaction costs given by a patent pool, Robert P. Merges and Michael Mattioli, ‘Measuring the Costs and Benefits of Patent Pools’ (2017) 78, 2 *Ohio State Law Journal*, 283 ff.

28 Giuseppe Colangelo, ‘Gli accordi di patent pooling’ (16 settembre 2008) *Società italiana di diritto ed economia* <<http://www.side-isle.it/ocs/viewabstract.php?id=141&cf=2>>.

29 *Ibid.*, 1.

30 Rai (n 9) 824.

31 Giuseppe Colangelo and Oscar Borgogno, ‘Data Sharing and Interoperability: Fostering Innovation and Competition through APIs’ (2019) 35, 5, *Computer Law & Security Review*, 105314, 105326.

32 *Ibid.*

- 25 The phenomenon of data pooling is being increasingly referred to by a strand of the literature with regards to the agreements made by firms for the sharing of “their digitalised information regarding a given market, in reference to a given service or generally in an industry, or within an e-ecosystem”³³. In this respect, data pools are complex collaborations that require collateral agreements on the processing technology needed for the pooling of the transferred data. The resulting agreements thus determine the processing infrastructure, which can be either delivered directly by one of the involved parties or outsourced by a third party³⁴.
- 26 With regards to the object of the transfer, the distinctive feature of data pooling practices is the difficulty to determine which data is exactly shared, this meaning the difficulty to determine whether only primary users’ data are being transferred or also the secondary data that are analytically drawn by the machine learning processes of the involved parties³⁵. In these regards, some strand of the literature³⁶ has interestingly observed that contracts regarding high technology projects “have become more and more fluid, because the projects are so complex that it is difficult to figure beforehand what is at stake”³⁷. This means, in turn that in the networked digital research environment, it is difficult to trace stable rules of data ownership and liability³⁸.
- 27 Under these premises, health data pools can be considered as a form of “contractually reconstructed research common”³⁹, which open up formed research data silos for the progression of scientific and technological progress⁴⁰. Hence, in the digital environment, the contractually-based aggregation of large health datasets owned by different research actors thus appears to serve innovation goals similar

33 Lindqvist (n[6]) 146.

34 Ibid.

35 Ibid, 149.

36 Karl Heinz Ladeur, ‘The Future of Law: Serial Law’ (2016) EUI Working Papers Law 2016/9 Department of Law <https://cadmus.eui.eu/bitstream/handle/1814/43345/LAW_2016_19.pdf?sequence=1&isAllowed=y>, 9.

37 Ibid.

38 Ibid, 6. This is very much observed by Effy Vayena and Alessandro Blasimme, ‘Health Research with Big Data: Time for Systemic Oversight’ (2018) 46 *The Journal of Law, Medicine & Ethics*, 119.

39 Reichman and Uhler (n 20) 416.

40 Mattioli (n 25) 187.

to the ones promoted by the patent system in a product-based economy. The contractual sharing of research valuable information is emerging as an increasingly important private ordering tool for the achievement of collaborative digital health innovation, in respect to which the intellectual property system alone appears to have a too little incentivising function⁴¹.

- 28 This raises in turn the issue whether at European policy and law, the sharing of health data between businesses for research and innovation purposes, is encouraged or rather restrained under different considerations as the ones related to the protection of health data subjects’ fundamental right, first of all to data protection. Against this backdrop, thus, the following paragraphs will assess whether and how health data sharing and the related innovation rationale is considered under European policy and the lawfulness of these data pooling practices under European data protection law.

D. Health Data Pools under European Policy: the Digital Single Market Strategy

- 29 Health data pools as described above involve i) massive processing of health data for the purposes of the delivery of digital health products and services and ii) the aggregation of different types of data among different stakeholders.
- 30 The first identified feature relates to the application of new processing infrastructures, such as algorithms and machine learning, for the purposes of the development of new tools and services based on information communication technologies (ICT). In this perspective, health data pools are to be inscribed in the broader economic phenomenon of digital health. In the words of the European Commission, “digital health and care refers to tools and services that use information and communication technologies (ICTs) to improve prevention, diagnosis, treatment, monitoring and management of health and lifestyle. Digital health and care has the potential to innovate and improve access to care, quality of care, and to increase the overall efficiency of the health sector”⁴².
- 31 From the second perspective, health data pools are to be placed in the other broader economic practice

41 Arguing in this sense Hilty (n 16) 898 ff.

42 For an overview European Commission, ‘eHealth: Digital Health and Care’ <https://ec.europa.eu/health/ehealth/overview_en>.

regarding information exchanges among different stakeholders. Information exchanges have been recently under increasing consideration by the European Commission, which has been stressing the importance of data sharing practices for the efficient development of the digital single market. In this context, the Commission has been employing the term “data sharing” in order to refer to “all possible forms and models” implying “data access or transfer” among different players, of both private and public nature⁴³. As the Commission further acknowledges, data sharing can be carried out through different technical mechanisms and under a variety of legal forms, supporting them⁴⁴. Under these premises, the practice of health data pools is to be contextualised in the two European policies regarding digital health and the free-flow of data. Far from being separate, these policies are intertwined fragments of the much wider European Digital Single Market Strategy.

I. Health Data Pools and Digital Health within the Digital Single Market Strategy

32 Digital health and the processing of health information have been increasingly considered at policy level by the European Commission for their innovation potential in the context of the digital internal market. This has ultimately led the Commission to comprehensively include digital health within the Digital Single Market Strategy for Europe⁴⁵. Hence, the digital transformation of European health and care can be considered in the general perspective of European digital markets.

- 33 Interestingly, the 2015 Digital Single Market Strategy for Europe⁴⁶ did not focus specifically on health and care, but already made some references to e-health. References to e-health were made as an example of another sector, amongst the others mentioned⁴⁷, where digital services would bring benefits to both users/consumers and businesses, particularly in terms of standardization and interoperability⁴⁸.
- 34 In May 2017, in the Communication on the Mid-Term Review on the implementation of the Digital Single Market Strategy, the European Commission came to strengthen the focus on digital health, particularly stressing the two policy objectives i) of providing citizens’ secure access to electronic health records and ii) of supporting data infrastructure to advance research, disease prevention and personalized health⁴⁹.
- 35 Ultimately, in its Communication on “enabling the digital transformation of health and care in the Digital Single Market: empowering citizens and building a healthier society”⁵⁰, the Commission has stressed the importance of the development of “strong approaches in high performance computing, data analytics and artificial intelligence, which can help design and test new healthcare products, provide faster diagnoses and better treatments”⁵¹.

43 So European Commission, ‘Commission Staff Working Document, Guidance on Sharing Private Sector Data in the European Data Economy, Accompanying the Document Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions “Towards a Common European Data Space’ (15 April 2018) SWD(2018) 125 final <<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018SC0125&from=EN>>, 5.

44 Ibid, 12.

45 See lately, European Commission, ‘Commission Communication on the Mid-Term Review on the implementation of the Digital Single Market Strategy. A Connected Digital Single Market for All’, (10 May 2017) COM(2017) 228 final; <<https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1496330315823&uri=CELEX:52017DC0228>>.

46 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 Digital Single Strategy for Europe, 2015 Digital Single Market Strategy for Europe’ (6 May 2015) COM(2015) 192 final <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2015%3A192%3AFIN>>.

47 E-Health has indeed been considered by the Commission together with other digital services in the context of e-government, e-energy-e-transport, *ibid*, 15.

48 European Commission, ‘Staff Working Document, Accompanying the Document- Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on Enabling the Digital Transformation of Health and Care in the Digital Single Market; Empowering Citizens and Building a Healthier Society’ (25 April 2018) <<https://ec.europa.eu/digital-single-market/en/news/communication-enabling-digital-transformation-health-and-care-digital-single-market-empowering>>, 3-4.

49 European Commission, (n 45) 19.

50 European Commission (n 48) 3.

51 *Ibid*.

- 36 According to the Commission, European health systems would benefit from digitization processes, in terms of resilience and sustainability⁵². Digital health tools are indeed deemed to improve patients' safety, reduce the number of avoidable mistakes, and improve the coordination and continuity of care and better adherence to treatment⁵³. These gains are evaluated within the frame of the resulting cost-savings and economic efficiencies⁵⁴.
- 37 The European Commission thus majorly links technological developments in health to the central goal of economic optimization and innovation⁵⁵. More precisely, the wider deployment of digital products and services in healthcare is deemed to stimulate growth and promote the European industry in the domain, with that overall maximizing the potential of the digital internal market⁵⁶.
- 38 Against the backdrop of the technological transformations relevant for the healthcare sector, the European Commission highlights the need for health and care authorities to face the emerging common challenges jointly. These challenges primarily concern the development of EU-wide standards for data quality, reliability and cybersecurity, the EU-wide standardization of electronic health records and a better interoperability through open exchange formats⁵⁷.

II. Health Data Pools and The Free Flow of Information Within the Digital Single Market Strategy

- 39 Health data pools are data sharing practices between different stakeholders, of both public and private nature, acting in the European internal market. From this perspective, health data pools are to be contextualised also in the other branch of European policy concerning the free-flow of information as lately concretised in the more specific policy promoting the accessibility and re-use of data.
- 40 Together with the rise of the digital economy, driven by “digital data, computation and automation”⁵⁸, the Commission has soon identified “the insufficient access to large datasets and the enabling infrastructure” as direct obstacles to market entry and to innovation⁵⁹. This is why the Digital Single Market Strategy has acknowledged information exchanges as a precondition for “maximising the growth potential of the digital economy” and assuring an efficient use of data across the EU⁶⁰.
- 41 Accordingly, the free-flow of information initiative⁶¹ has become a key action within the project of the implementation of a Digital Single Market Strategy⁶². In particular, the importance of access to health data has been lately highlighted by the European

52 Ibid.

53 Ibid, 11.

54 Ibid, 1; 11 and 12. The market efficiency gains of digitisation of healthcare have been stressed by the Council of Europe on several occasions. See Council of the European Union, ‘Council conclusions: Towards modern, responsive and sustainable health systems’ (6 June 2011) OJ C 202, 8 July 2011, 10; Id., ‘Council conclusions on the “Reflection process on modern, responsive and sustainable health systems’ (10 December 2013) OJ C 376 21 December 2014, 3; Id., ‘Council Conclusions on the Economic Crisis and Healthcare’ (20 June 2014) OJ C 217 10 July 2014, 2; Id., ‘Council Conclusions on Personalised Medicine for Patients’, 7 December 2015, OJ C 421 17 December 2015, 2. Id., ‘Council Conclusions on Health in the Digital Society- Making Progress in Data-driven Innovation in the Field of Health’ (2017) OJ C 440/3 21 December 2017, 5.

55 Mark L. Flear, ‘Regulating New Technologies: EU Internal Market Law, Risk and Socio-Technical Order’ in: Marise Cremona, *New Technologies and EU Law* (Oxford University Press, 2017) 74 ff., 76.

56 European Commission (n 48) 5.

57 Ibid, 5.

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- Towards a Thriving Data Economy’ (2 July 2014) COM(2014) 442 final <<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52014DC0442&from=EN>> 2.

59 Ibid, 2-3.

60 European Commission (n 46), 14-15.

61 The free flow of information initiative was first announced in the “Mid-Term Review on the implementation of the Digital Single Market Strategy”. See also European Commission, ‘Commission Staff Working Document on the Free Flow of Data and Emerging Issues of the European Data Economy’ (10 January 2017) SWD(2017) 2 final <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52017SC0002>>, 30-31.

62 European Commission, ‘Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions-Building a European Data Economy’ (10 January 2017) COM(2017) 9 final <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2017:9:FIN>>.

Commission in the “European strategy for data”⁶³. Here, the establishment of a “common European health data space” has been considered among the nine European data spaces the European Commission intends to encourage through the newly established strategy⁶⁴. For the purposes of strengthening the relevant regulatory framework, the Commission has announced a new package of measures, meant to create a European common data space, in which new products and services are developed upon the shared data⁶⁵.

- 42 In this respect, the Commission has come to stress the relevance of privately held data for the purposes of business to business (B2B) sharing agreements⁶⁶. It is highlighted that access and use of a same set of shared data can be employed by businesses for the development and the testing of different products⁶⁷.
- 43 In addition to this, also data transfers occurring within public-private partnerships have been considered by the Commission for their economic potential⁶⁸. In this perspective, it is interesting to highlight that the reform of the Public Sector Information Directive places a particular emphasis on research data⁶⁹. In this respect, the new Open Data Directive⁷⁰ expressly considers research data under art. 10 stating that “member states shall support the availability of research data (...)” on the basis of “open access policies”. Access to and reuse of publicly funded research data is further encouraged by the renewed Recommendation on access to

and preservation of scientific information⁷¹. The Recommendation considers the new text and data mining technologies⁷² and the technical standards for data⁷³ as important catalysts for the access and reuse of extracted scientific information generated by public stakeholders. Accordingly, the new Recommendation on access to and preservation of scientific information adapts these goals to the new datification courses and the enhanced data analytics capabilities⁷⁴. Big data are indeed deemed to change the way research is performed and knowledge is shared⁷⁵, along the lines of a paradigm shift towards more collaborative methods of carrying out scientific research⁷⁶. This is in turn leading to a more open and transparent research approach, which in the view of the Commission needs to be further encouraged and incentivised⁷⁷.

- 44 Both the new Open Data Directive and the mentioned Recommendation appear to directly build upon the “principle of free movement of data within the EU”⁷⁸, in this way complementing the Regulation regarding the free-flow of non-personal data⁷⁹.
- 45 Against the backdrop of these first legislative measures regarding the free flow of data within the Digital Single Market, the question has arisen in the literature whether the European policy regarding

63 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.

64 Ibid, 22.

65 European Commission (n 43) 1.

66 Ibid, 5.

67 Ibid, 2.

68 European Commission, ‘Big Data Value Private-Public Partnership’ <<https://ec.europa.eu/digital-single-market/en/big-data-value-public-private-partnership>>.

69 European Commission (n 43) 6-7.

70 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 (26 June 2019), OJ L 172/56 <<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L1024&from=EN>>.

71 European Commission, ‘Commission Recommendation EU 2018/790 of 25 April 2018 on Access to and Preservation of Scientific Information’ <https://www.eoscportal.eu/sites/default/files/CELEX_32018H0790_EN_TXT.pdf>.

72 Ibid, para. 3, titled “Management of Research Data, including Open Access”.

73 Ibid, para 6 and 7, titled “Infrastructures for Open Data”.

74 Ibid, recital 12.

75 Ibid, recital 2.

76 Ibid, recital 9, stressing that “technological progress has over time caused a major shift in the world of science towards increasingly collaborative methods, and has steadily contributed to an increasing volume of scientific material”.

77 Ibid, recital 10 and para 9, titled “Incentives and Rewards”.

78 European Commission (n 43) 10.

79 European Parliament and Council of the European Union, ‘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’ 28 November 2018, OJ L 303/59, online available at <<https://eurlex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32018R1807&from=EN>>.

the free flow of information should concern only non-personal data or include also personal data. Indeed, it was initially declared that the free-flow of information would have referred only to non-personal data⁸⁰. Personal data were said to fall outside the scope of the free-flow of data initiative since this data is already regulated in the different regulatory sector covered by the General Data Protection Regulation and the e-Privacy Directive, specifically setting the framework with respect to processing of personal data⁸¹.

- 46 However, personal data have been somehow taken into consideration by the Commission, acknowledging that actors in the data economy “deal both with personal and non-personal data and that data flows and datasets will regularly contain both types”⁸². It is also further affirmed that “any policy measure must take account of this economic reality and of the legal framework on the protection of personal data, while respecting the fundamental rights of individuals”⁸³. These words by the Commission reflect that the object of the policy regarding the free-flow of information is still largely unclear⁸⁴. This is highlighted by a strand of the literature, calling for a more comprehensive policy and regulatory approach⁸⁵. Along these lines, the European Commission has lately come to pair the General Data Protection Regulation with the Regulation on the free flow of non-personal data, considering the two bodies of law as a comprehensive and coherent framework to the free movement of data in the European Union⁸⁶.

80 European Commission (n 43) 1.

81 Ibid.

82 European Commission (n 62) 9.

83 Ibid.

84 Noticing a certain ambivalence by the Commission with regards the relationship between the free-flow of data policy and data protection law, Inge Graef, Raphaël Gellert and Martin Husovec, ‘Towards a Holistic Regulatory Approach for the European Data Economy: Why the Illusive Notion of Non-Personal Data is Counterproductive to Data Innovation’ (2019) 44, 5 *European Law Review* 605, 607.

85 Ibid., 610; Josef Drexl, ‘Legal Challenges of the Changing Role of Personal and Non-personal Data in the Data Economy’ in: Alberto De Franceschi and Reiner Schulze (ed.), *Digital Revolution- New Challenges for Law- Data Protection, Artificial Intelligence, Smart Products, Blockchain Technology and Virtual Currencies* (C.H. Beck, 2019) 19, 23 ff.

86 European Commission, ‘Free Flow of Non Personal Data’ <<https://ec.europa.eu/digital-single-market/en/free-flow-non-personal-data>>. See in these regards, Drexl (n 85) 20,

E. Health Data Pools as Health Data Processing under European Data Protection Law

- 47 Health data pools for research and innovation purposes in the field of digital health involve the sharing and thus processing of data subjects’ actual or potential sensitive information. The innovation objectives underlying health data pools and supported by European Union’s policy in the context of Digital Single Market Strategy thus need to be weighed against other regulatory objectives of European law and especially of European data protection law.
- 48 The General Data Protection Regulation sets a specific regulatory framework for the processing of health data. Indeed, it provides specific definitions of different types of health data, such as genetic data or biometric data under art. 4(13, 14) and 15 GDPR. In addition, it categorizes health data as a “special category of data” the processing of which is prohibited under art. 9(1) GDPR. Ultimately, it sets some broad exemptions to such prohibition. These exemptions allow the processing of health data if it is carried out for certain purposes and provided specific conditions are met.
- 49 By establishing a general prohibition of health data processing and some grounds of exceptions to that prohibition, the regulatory status of health data processing under the GDPR appears to be defined by a layered regime and triggers some challenging interpretative efforts.
- 50 Before digging deeper into the multifaceted data protection law provisions regarding the processing of health data, some theoretical background considerations are needed. Indeed, the layered regime established with regards to health data is the result of a much deeper tension within European data protection law, which the General Data Protection Regulation has inherited from the previous Directive and partly exacerbated. This tension relates to the two seemingly contrasting objectives of data protection law, on the one hand the protection of data subjects’ fundamental rights in the digital environment and on the other hand the promotion of lawful data flows fueling efficiency outcomes within the digital single market.

observing that “personal data are no longer only objects of a privacy interest but are increasingly recognised in their role as a valuable asset used by businesses in the digital sector”.

I. European Data Protection Law between Fundamental Rights Protection and Market Regulation

- 51 Born from the rib of the right to privacy⁸⁷, the European right to data protection has become an autonomous fundamental right in the European Charter of Fundamental Rights under art. 8 EU Charter⁸⁸. This is directly reflected in the General Data Protection Regulation⁸⁹, which is legally rooted in art. 16 TFUE and recalls art. 8 EU Charter in recital 1.
- 52 The fundamental rights dimension of the European right to data protection has however broadened in the digital economy, where data processing activities pose substantial threats first of all to individuals' rights to autonomy and informational self-determination⁹⁰, and also to other fundamental

rights, such as the right to informational self-determination, the right to equality and non-discrimination⁹¹.

- 53 As a direct response to the ongoing technological and economic changes, the General Data Protection Regulation follows a risk-based approach, which considers the treatment of personal data conducted on a massive scale⁹² as a risky practice⁹³. From this perspective, the protection of the right to data protection in the form of the right to a fair, transparent and accountable data collection and processing⁹⁴ becomes a structural precondition to the protection of these other fundamental rights, as jeopardised by businesses' algorithmic models⁹⁵.
- 54 However, the General Data Protection Regulation's objective of protecting data subjects' fundamental rights from the intrusiveness of new data processing technologies⁹⁶ coexists with a further regulatory

87 For a comment on the relationship between privacy and data protection, Raphaël Gellert and Serge Gutwirth, 'The Legal Construction of Privacy and Data Protection' (2013) 29 Computer Law & Security Review, 522 ff.; Orla Lynskey, 'Deconstructing Data Protection: the 'Added-value' of a Right to Data Protection in the EU Legal Order' (2014) 63 International and Comparative Law Quarterly, 569 ff.

88 For a critical assessment of the fundamental rights nature of the right to data protection, see Bart Van Der Sloot, 'Legal Fundamentalism: is Data Protection Really a Fundamental Right?' in: Ronald Leenes, Rosamunda van Brakel, Serge Gutwirth and Paul De Hert, Data Protection and Privacy: (In) visibilities and Infrastructure (Springer, 2017), 3 ff.

89 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), from now on GDPR.

90 Alessandro Spina, 'Risk Regulation of Big Data: Has the Time Arrived for a Paradigm Shift in Eu Data Protection Law?', Case notes to Case C-293/12 and C-594/12 *Digital Rights Ireland and Seitlinger and Others*' (2014) 5, 2 European Journal of Risk Regulation, 248 ff., 251, commenting on the statements of the European Court of Justice, affirming that the various collected "(...) data, taken as a whole may allow very precise conclusions to be drawn concerning the private lives of the persons whose data has been retained". So Court of Justice of the European Union, *Digital Rights Ireland Ltd vs. Seitlinger and Others* (8 April 2014) Joined Cases C293/12 and C594/12, <<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62012CJ0293&from=EN>>, para 27.

91 See Recital 75 GDPR: "the risk to the rights and freedoms of natural persons, of varying likelihood and severity, may result from personal data processing which could lead to physical, material or non-material damage (...)". Emphasis added. Sandra Wachter, 'Primus inter Pares: Privacy as a Precondition for Self-development, Personal Fulfilment and the Free Enjoyment of Fundamental Rights' (22 January 2017) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2903514&download=yes>.

92 See Recital 6 GDPR observing how "rapid technological developments and globalisation have brought new challenges for the protection of personal data. The scale of the collection and sharing of personal data has increased significantly. Technology allows both private companies and public authorities to make use of personal data on an unprecedented scale in order to pursue their activities".

93 See Recitals 75-76 GDPR. For the literature see Ira S. Rubinstein, *Big Data: The End of Privacy or a New Beginning?* (2013) 3, 2 International Data Privacy Law, 74 ff., highlighting the systemic risks related to massive data processing and Raphaël Gellert, 'Understanding Data Protection as Risk Regulation' (2015) Journal of Internet Law, 3, 6 ff.

94 See art. 5 GDPR.

95 Viktor Mayer-Schonberger-Kenneth Cukier, *Big Data: A Revolution that Will Transform How We Live, Work and Think* (Houghton Mifflin, 2013), 20, noticing that data protection was generated as a risk regulation, aimed at controlling the different steps of data processing operations, made up by "complex and rich procedures to control and regulate the use of technology".

96 For a critical of the GDPR in respect to algorithmic inferences, Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law

pillar of European data protection law, related to the promotion of the free flow of personal information for the integration and consolidation of the internal market. This pillar had a primary importance within the Data Protection Directive⁹⁷, whose legal foundations were to be found exactly in the regulation of the internal market under art. 100a of the Treaty establishing the European Community⁹⁸. It has however not lost its hold within the normative system of the General Data Protection Regulation. As has been observed by prominent scholarship, under the new Regulation the fundamental rights and the market integration purposes appear to be placed “on equal footing”⁹⁹.

- 55 Here, the market integration objective comes right behind the primary objective of data subjects’ fundamental rights in the digital economy, and is expressed in recital 2 GDPR, stating how the Regulation is intended to contribute amongst others, “to the economic and social progress” and “to the strengthening and the convergence of the economies within the internal market”. Accordingly, recital 5 GDPR acknowledges how the flows of personal data have increased as a consequence of the “economic and social integration resulting from the functioning of the internal market” and with that also the “exchange of personal data between public and private actors”. This is confirmed also by recital 13 GDPR, where the free movement of personal data is considered as a requirement for the proper functioning of the internal market and ultimately by recital 123 GDPR, where supervisory authorities are given the task of monitoring and contributing to the application of data protection rules “in order to

protect natural persons in relation to the processing of their personal data and to facilitate the free flow of personal data within the internal market”¹⁰⁰.

- 56 These statements reflect the acknowledgment by the European legislator of the economic value of personal data within the dynamics of the digital economy. They reflect the view that personal data -and the sharing of it- are not only an object of protection but also an “innovation enabling technology”¹⁰¹ and with that a strategic asset for the establishment of an efficient Digital Single Market¹⁰².
- 57 Against the backdrop of the cited recitals, it appears that under the Regulation more than it occurred in the Directive, European data protection law is characterised by an internal tension between two apparently conflicting aims, on the one hand the restriction of personal data processing for the sake of the protection of the data subjects’ rights and on the other hand the maximisation of personal data flows for the development of the digital economy¹⁰³.

II. The Legal bases for the Processing of Health Data

- 58 The two above-highlighted objectives of European data protection law are well reflected in the regulation of health data established by the General Data Protection Regulation.
- 59 Indeed, in line with the previous Data Protection Directive¹⁰⁴, the General Data Protection Regulation

in the Age of Big Data and AI’ (2019) 2 Columbia Business Law Review, 494 ff.

- 97 See in these regards also the European Court of Justice, ‘Österreichischer Rundfunk and Others’ (20 May 2003) Joined Cases C-465/00, C-138/01 and C-139/01 <<http://curia.europa.eu/juris/showPdf.jsf?text=&docid=48330&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=8237402>>, para 39, and Id., ‘Commission v. Germany’ (9 March 2010) Case C-518/07 <<http://curia.europa.eu/juris/document/document.jsf?text=&docid=79752&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&id=8240530>>, para 3.
- 98 See art. 100 Treaty Establishing the European Community (N C 224/6 OJ 31 August 1992) <<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:11992E/TXT&from=EN>>. For the literature see Van Der Sloot (n[88]), 25.
- 99 Orla Lynskey, *The Foundations of EU Data Protection Law* (Oxford University Press, 2015) 47.

100 So Recital 123 GDPR.

101 Urs Gasser, ‘Cloud Innovation and the Law: Issues, Approaches and Interplay’ (18 March 2014) Berkman Center Research Publication, 2014-7 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2410271>, 6 ff.

102 Luca Marelli and Giuseppe Testa, ‘Scrutinizing the EU General Data Protection Regulation- How Will New Decentralized Governance Impact Research?’ (4 May 2018) 360, 6388 Science, 496, 497-498.

103 For a reconstruction of the “hybrid nature of EU data protection law”, Lynskey (n 99) 8-9.

104 See art. 8 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ L 281, 23 November 1995. European Data Protection Supervisor, ‘Opinion on the Communication from the Commission on ‘eHealth Action Plan 2012-2020- Innovative Healthcare for the 21st Century’ (27 March 2013) <https://edps.europa.eu/sites/edp/files/publication/13-03-27_ehealth_

subjects the processing of health data to stricter data protection rules. The prohibition of processing special categories of data, under art. 9(1) GDPR constitutes a direct (over-)regulatory response to the objective of protecting data subjects' fundamental rights against non-consented accesses to very intimate subjective spheres such as the one of health¹⁰⁵.

- 60 However, there are some exceptions to this prohibition, which allow the processing of health data on the basis of different legal grounds listed under art. 9(2) GDPR¹⁰⁶.
- 61 These legal grounds can be respectively sub-grouped as follows: i) data subject's consent under art. 9(2) lett. a) GDPR and, strictly related to it, the need to protect a vital interests of the data subject under art. 9(2) lett. c) GDPR as well as the manifest publicity of the personal data under art. 9(2) lett. e) GDPR; ii) the processing is necessary for reasons of substantial public interest under art. 9(2) lett. g) GDPR, for the purposes of preventive or occupational medicine, medical diagnosis, the provision of health or social care or treatment or the management of health or social care and systems and services under art. 9(2) lett. h) and for reasons of public interest in the area of public health under art. 9(2) lett. i) GDPR; iii) the processing is necessary for scientific or historical research purposes or statistical purposes under art. 9(2) lett. j) GDPR.
- 62 The first category of legal bases for the processing of health data is based on the data subjects' subjective perspective, concretised through his/her determinations in the form of consent or in respect to his/her fundamental interests. Conversely, the other two identified categories take a rather objective perspective and rely on objective features of data controllers' processing activities, related to their public interest or research-oriented nature¹⁰⁷.

action_en.pdf>, 3 and Article 29 Data Protection Working Party, 'Working Document on the Processing of Personal Data Relating to Health in Electronic Health Records' (15 February 2007) <<https://www.dataprotection.ro/servlet/ViewDocument?id=228>>, 8.

- 105 Stressing the symbolic value of this provision Tal Z. Zarsky, 'Incompatible: The GDPR in the Age of Big Data' (2017) 47 *Seton Hall Law Review* 995, 1014.
- 106 Nicolo Zingales, 'Data Protection Considerations in EU Competition Law: Funnel or Straightjacket for Innovation?', in: Paul Nihoul and Pieter Van Cleynenbreugel (ed.), *The Roles of Innovation in Competition Law Analysis* (Edward Elgar, 2018) 79 ff., 108, considering data protection law as a "permission based" regime.
- 107 In this direction see, Marelli and Testa (n 102) 496, observing a "shift toward a decentralized, controller-anchored, and

- 63 As a general premise it needs to be recalled that the mentioned legal bases established under art. 9(2) GDPR for the processing of special categories of data need to be linked to the legal grounds generally established under art 6 GDPR setting the conditions for the lawfulness of the processing. According to the majority of the scholarship indeed, the legal grounds under art. 9(2) GDPR are complementary to the general requirements for a lawful data processing under art. 6 GDPR. This means that the existence of a general lawful basis under art. 6 GDPR is a precondition for the processing of special categories of data under the special conditions laid down under art. 9(2) para GDPR¹⁰⁸.
- 64 As will be better shown in the next paragraph, the legitimate basis for processing under art. 9(2) lett. j) GDPR appears to be particularly interesting for the case of health data pools. It indeed appears to provide some fertile normative grounds for the flourishing of health data pools aimed at developing and placing new digital health products and services on the market. By doing so, it attests the European legislator's acknowledgement of the scientific- and thus of the innovation- enabling value of health data as special categories of data within the European digital market.
- 65 This legal basis for the processing of health data needs to be carefully interpreted in respect to the general prohibition regarding the same processing of special categories of personal data. As will be shown, it is also connected to an outright "research exemption", derogating to important general data protection principles and rules. If correctly implemented, this exemption does not totally back out fundamental rights protection goals. However, as will be argued, due to the interpretative uncertainties that it raises, it opens some loopholes that risk doing so.

III. Research as a Legal Basis for the Processing of Health Data

- 66 Among the above-mentioned legal bases for the processing of health data, the most interesting one for the case of health data pools is given by art. 9(2) lett. j) GDPR. This provision allows health
-
- accountability-based model".
- 108 This is the solution given by Edward S. Dove, 'The EU General Data Protection Regulation: Implications for International Scientific Research in the Digital Era' (2018) *The Journal of Law, Medicine & Ethics*, 1013, 1024. See also Sebastian Schulz, 'Art. 9 Verarbeitung besonderer Kategorien personenbezogener Daten' in: Peter Gola, *Datenschutz-Grundverordnung VO (EU) 2016/679- Kommentar* (C.H. Beck, 2018), ed.) 361 ff., 365.

data processing when it is “necessary for reasons of public interest, scientific or historical research purposes or statistical purposes”. In this perspective, thus, art. 9(2) lett. j GDPR establishes an autonomous legitimate basis for the processing of health data, which is directly grounded in research objectives.

- 67 The promises of health data processing for scientific research projects is acknowledged under recital 157 GDPR, where it is stated that “by coupling information from registries researchers can obtain new knowledge of great value with regard to widespread medical conditions such as cardiovascular disease, cancer and depression. (...) In order to facilitate scientific research personal data can be processed for scientific research purposes, subject to appropriate conditions and safeguards set out in Union or Member State law”.
- 68 In accordance with these statements, processing for research purposes appears to have a privileged position within the General Data Protection Regulation, which provides various definitions of data-driven research. The Recitals do in fact treat different types of research separately, distinguishing between “scientific research”, “historical research”, “statistical research”.
- 69 With regards to scientific research, recital 159 GDPR defines it as “the technological development and demonstration, fundamental research, applied research, and *privately funded research*”¹⁰⁹, as well as public health research. The recital expressly refers to Article 179(1) of the Treaty on the Functioning of the European Union, which encourages “the objective of strengthening its scientific and technological bases by achieving a European research area in which researchers, scientific knowledge and technology circulate freely”. As clarified by recital 160 GDPR, historical research comprises genealogical research. Ultimately, “statistical research” is defined under recital 162 GDPR, as “any operation of collection and the processing of personal data necessary for statistical surveys or for the production of statistical results”. As the same recital affirms, statistical research “implies that the result of processing for statistical purposes is not personal data, but aggregate data”. While statistical research may be used in support of scientific research, it cannot be “used in support of measures or decisions regarding any particular natural person”¹¹⁰.

109 Emphasis added.

110 The Recital specifies that the EU or the Member States should legislate around the scope of the statistical research exemptions, including defining the appropriate safeguards for assuring “statistical confidentiality”. So recital 162 GDPR.

- 70 A strand of the literature commenting art. 9(2) lett. j) GDPR, has observed that the notion of processing for statistical purposes could encompass also processing activities carried out through big data analytics as they rely exactly on statistical methods¹¹¹. As can be derived from the mentioned recitals, the General Data Protection Regulation, adopts a broad definition of research¹¹², likely to encompass the activities of both public and private entities¹¹³. These considerations lead to the question of the nature of the link between the legal grounds of processing for research purposes and for public interest.
- 71 Indeed, although it is true that art. 9 (2) lett. j) GDPR refers both to processing activities carried out in the public interest and for research purposes, the notions are considered in a separate manner by the Regulation¹¹⁴. By considering the research purpose autonomously, indeed, the Regulation appears to overcome the approach adopted by the previous Directive, which mentioned the scientific research as an example of “reasons of substantial public interest” under recital 34¹¹⁵. It thus seems that, differently from what was the case under the Directive, under the Regulation scientific research is not a specification of the public interest.
- 72 In view of the risk of reliance on the legal grounds of scientific research also for commercially-oriented activities¹¹⁶, the Biobanking and

111 Wachter and Mittelstadt (n 96) 592-; similarly Zarsky (n 105) 1013.

112 This is directly affirmed by recital 159 GDPR, which affirms that “for the purposes of this Regulation, the processing of personal data for scientific research purposes should be interpreted in a broad manner”.

113 Similarly, Kärt Pormeister, ‘Genetic Data and the Research Exemption: is the GDPR Going too Far?’ (2017) 7, 2 International Data Privacy Law, 137 ff.

114 Paul Quinn and Liam Quinn, ‘Big Genetic Data and Its Big Data Protection Challenges’ (2018) Computer Law & Security Review, 1015.

115 Mahsa Shabani and Pascal Borry, ‘Rules for Processing Genetic Data for Research Purposes in View of the New General Data Protection Regulation’ (2018) 26, 2 European Journal of Human Genetics, 149, 153. It must be additionally recalled that under the Previous Directive, the legal base of the processing in the public interest, has been used by Member States to permit processing for a range of purposes, as scientific research. This has occurred for example in Germany. See Quinn and Quinn (n 114) 1013.

116 Chih-hsing Ho, ‘Challenges of the EU General Data Protection Regulation for Biobanking and Scientific Research’ (2017) 25, 1 Journal of Law, Information and Science, 84, 98-99,

BioMolecular Resources Research Infrastructure-European Research Infrastructure Consortium (BBMRI-ERIC) has stressed the need to restrict the broad interpretation given to the General Data Protection Regulation's notion of scientific research so as to consider only public interest-oriented research activities¹¹⁷. A first restriction for these purposes is directly provided under art. 9(2) lett. j GDPR, requiring processing activities carried out for research purposes to be based on Union or Member State law. This means that the well before interpretative debates, the definition of which processing activities shall fall under art. 9(2) lett. j) GDPR is left to specific legislations under Union or Member State law. With regards to European Union law, an example of such specific regulation is given by the Clinical Trial Regulation¹¹⁸, which the European Data Protection Board has lately clarified as a "sectoral law containing specific provisions relevant from a data protection viewpoint but no derogations to the GDPR", thus clarifying that the two frameworks both apply simultaneously¹¹⁹.

where the Author cites some empirical studies showing the mistrust of consumers with regards the use of health data by private commercial entities. See Royal Statistical Society, 'Royal Statistical Society Research on Trust in Data and Attitudes Toward Data Use/Data Sharing-Briefing Note' (22 July 2014) <<http://www.statslife.org.uk/images/pdf/rss-data-trust-data-sharingattitudes-research-note.pdf>>.

- 117 BBMRI-ERIC- Biobanking and Biomolecular Resources Research Infrastructure, 'Position Paper on the General Data Protection Regulation' (October 2015) <http://www.bbMRI-eric.eu/wp-content/uploads/BBMRI-ERIC-Position-Paper-General-Data-Protection-Regulation-October-2015_rev1_title.pdf>, 3. This is the view shared also by a strand of the literature, Bertram Raum, 'DS-GVO Art. 89 Verarbeitung zu Archivzwecken, Forschungszwecken' in: Eugen Ehmann and Martin Selmayr (ed.), *Datenschutz-Grundverordnung* (C.H. Beck, 2017), 41-42 and William Nicholson Price II, Margot E. Kaminski, Timo Minssen and Kayte Spector-Bagdady, 'Shadow Health Records Meet New Privacy Laws- How Will Research Respond to a Changing Regulatory Space?' (2019) 363, 6426 Science, 448, 450.
- 118 Regulation EU n. 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, OJ L 158/1, 27 May 2014 <https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/reg_2014_536/reg_2014_536_en.pdf>.
- 119 European Data Protection Board, 'Opinion 3/2019 Concerning the Questions and the Answers on the Interplay Between the Clinical Trials Regulation (CTR) and the General Data Protection Regulation (GDPR)(Art. 70.1.b)' (32 January 2019) <<https://www.dataprotection.ro/servlet/ViewDocument?id=1629>>, 3.

73 Under these premises, it appears that the General Data Protection Regulation leaves much room open for interpretation regarding the link between the processing for research- be it scientific or statistical- purposes and secondary commercially-oriented purposes. In this respect, however, the same art. 9(2) lett. j) GDPR sets some first normative limits for the processing of health data for research purposes, requiring such processing to be proportionate to the aim pursued- consistently with the proportionality and data minimization principles under art. 5(1) lett. b) GDPR-, to respect the essence of the data protection right and be subject to specific safeguards for the protection of the data subjects' fundamental rights and interests¹²⁰. Hence, in addition to further legislative definitions, more specific and decisive interpretative guidelines from the European Data Protection Board regarding such limits would be desirable¹²¹.

IV. The Special Data Protection Regime for the Processing of Health Data under the Research Exemption

- 74 A correct interpretation of the scope of art. 9(2) lett. j) GDPR is of crucial importance in order to determine the severity of the data protection regime applicable to the case of health data pools. In the General Data Protection Regulation's system, the processing of personal data for research purposes is indeed related to a special data protection regime, which entails significant derogations to ordinary data subjects' rights and controllers' obligations and at the same time requires the enactment of adequate safeguards for the protection of data subjects' rights in the context of data-driven research projects.
- 75 Such special data protection regime is given by the interplay between the considered art. 9(2) lett. j) GDPR and arts. 5(1) lett. b); 6(4); and 89 GDPR. The interaction between the cited provisions subjects also data concerning health, which are processed under the legitimate basis set out under art. 9(2) lett. j) GDPR, to the "research exemption" established under arts. 5(1) lett. b); 6(4); and 89 GDPR. These last provisions state that further processing of personal data for research purposes is *per se* compatible with the initial purpose of data collection, provided

120 Giovanni Comandè, 'Ricerca in sanità e data protection... un puzzle risolvibile' (2019) 1 Rivista italiana di medicina legale, 187, 195.

121 The need for a clarification regarding the scope of the GDPR's research exemption is stressed by Price, Kaminski, Minssen and Spector-Bagdady (n 117) 450.

the safeguards required under art. 89(1) GDPR are enacted. Accordingly, under the research exemption, the processing of health data for research purposes can derogate fundamental data protection principles, such as the principle of purpose limitation under art. 5(1) lett. b) GDPR. Likewise, the principle of storage limitation under art. 5 (1) lett. e) GDPR can be subject to derogations in case personal data are processed for research purposes. As a result, if necessary for research purposes, health data may be stored for longer periods and be employed for wider purposes than would be otherwise allowed under the general data minimization principle¹²².

- 76 Also data subjects' rights as the right to be forgotten under art. 17(3) GDPR and the right to be informed under art. 14(5) lett. b) GDPR can be derogated in case the enactment of the right impairs the achievement of the research objectives¹²³. However, controllers' information duties under art. 13 GDPR remain effective in case the data used for research purposes is directly collected from data subjects, unless, as specified by recital 62 GDPR, "the provision of information to the data subject proves to be impossible or would involve a disproportionate effort".
- 77 The compression of the information controllers have to disclose in the context of research projects sensitively weakens data subjects' control prerogatives over their health data, which under the ordinary data protection regime are addressed by controllers' transparency obligations especially in the privacy notice under art. 14(1) GDPR¹²⁴. The mentioned derogations to controllers' ordinary obligations well reflect the controller-oriented nature of research as a legal basis for processing. These derogations indeed allow the data controller to take full control over the data analysed for research purposes. This transfers the control barycenter onto the processing entities, without the data subjects knowing the conditions under which their health personal data are processed¹²⁵.

122 See art. 5(1) lett. c) GDPR.

123 As observed by some scholars, compliance with the transparency requirements within long data-driven research projects could be disproportionate and substantially impair the objectives of the processing, especially when there are many data subjects involved and the data has been heavily pseudonymised. So Quinn and Quinn (n 114) 1014.

124 Dove (n 108) 1024.

125 Pormeister (n 113)139, observing that "the exceptions from the storage and purpose limitations afforded to the research exemption create an outcome in which consent will become more irrelevant over time in correlation with

78 Additional derogations from the ordinary data protection regime set out by the Regulation can be further provided by Member State law: art. 89(2) GDPR enables Union or Member State law to provide derogations from data subjects' right to access under art. 15 GDPR; right to rectification under art. 16 GDPR; right to restriction of processing under art. 18 GDPR and ultimately the right to object under art. 21 GDPR¹²⁶. Under art.89(2) GDPR, controllers can derogate to these rights when these "are likely to render impossible or seriously impair the achievement of the specific purposes" and the derogations are necessary for the fulfilment of the purpose¹²⁷.

79 In order to counterbalance of these derogations, art. 89(1) GDPR conditions the processing of personal data for research purposes to the enactment of appropriate "technical and organizational measures" needed in order to safeguard "the rights and freedoms of the data subject". A first relevant safeguard is directly mentioned by art. 89(1) GDPR, which refers to pseudonymization of research data.

80 Art. 9(4) GDPR leaves however the definition of such safeguards to Member States' discretion in establishing "further conditions, including limitations, with regard to the processing of genetic data, biometric data or data concerning health". In this perspective, codes of conduct, whose enactment is recommended under art. 40 GDPR could be relevant tools for the establishment of further data protection safeguards for health research. Accordingly, the ultimate degree of the restrictions posed by data

advancements in personal medicine".

126 It must be observed that the possibility granted to national legislations to derogate from the right to object under art. 21 GDPR expressly recalled by art. 89, 2 para GDPR, is to be reconciled with the provision under the same art. 21, 6 para GDPR, affirming the endurance of the right at stake in case of processing carried out for "scientific or historical research purposes or statistical purposes pursuant to art. 89, 1 GDPR". As can be derived from art. 21, 6 para GDPR, derogation to the data subjects' right to object is admitted when "the processing is necessary for the performance of a task carried out for reasons of public interest". This is thus the rule in absence of any national legislation. Conversely, a national legislation can under art. 89, 2 para GDPR derogate to the rule in case the exercise of the right is likely to render impossible or seriously impair the achievement of the specific (research) purposes and in case the restrictions are necessary to fulfil the purpose. Dove (n 108) 1025.

127 Art. 89,2 para GDPR. With regards to processing for scientific purposes, the English Data Protection Bill approved in 2018, has established derogations with regards to the right to access under art. 15 GDPR; to rectification under art. 16 GDPR; to object under art. 21 GDPR.

protection law to the processing of health data will largely depend on how burdensome the conditions and safeguards defined at national level or in codes of conduct will be¹²⁸.

- 81** In the absence of these national determinations, the special regime set by the General Data Protection Regulation for research activities, establishing the above-mentioned derogations to ordinary principles and rules, is directly applicable. The “relaxation of the law” resulting from the traced special data protection regime related to research thus enables businesses to share and thus process health data in the context of digital health-related research projects. This means that the General Data Protection Regulation ultimately appears to encourage health data pools established for research and innovation purposes, rather than curbing them.
- 82** The underlying risk of such special data protection regime is that big data controllers that participate to health data pools end up creating new statistical models based upon users’ special categories of data. These models could in turn facilitate “discrimination by association”¹²⁹ strategies in the broader digital market¹³⁰. In view of the derogations to data subjects’ rights under the data protection regime for research, data subjects would have weaker reaction means with regards to the results of these statistical enquiries¹³¹.
- 83** In this respect, it needs however to be recalled that also in respect to the processing of health data for research purposes, important data subjects’ rights are still applicable. In this perspective, reference needs to be made, in particular, to the right not to be subject to automated decisions under art. 22 GDPR. This right is specifically taken into consideration under the already mentioned recital 162 GDPR, which prohibits the use of personal data in the context of research activities “in support

of measures or decisions regarding any particular natural person”¹³².

- 84** As the recital suggests, thus, processing of personal data carried out for research purposes cannot result in profiling activities and other decisions regarding single natural persons¹³³. This statement is extremely important and poses some interesting normative grounds for interpreting the special data protection regime regarding data-driven research in a way that prevents research processing activities over health data from triggering further, “secondary” commercial actions.
- 85** First solutions in this respect could be found in the realignment of the notion of research relevant under data protection law to public interest-oriented processing purposes. This would imply the re-application of the “full” ordinary data protection law regime, in case health data are further used for commercial purposes, that is, for the commercial employment of the statistical models designed in the context of research projects¹³⁴.

128 Paul Quinn, ‘The Anonymisation of Research Data- a Pyrrhic Victory for Privacy that Should not be Pushed too Hard by the EU Data Protection Framework?’ (2016) 24 *European Journal of Health Law*, 1–21.

129 This term is used by Sandra Wachter, ‘Affinity Profiling and Discrimination by Association in Online Behavioural Advertising’ (2020) 35, 2 *Berkeley Technology Law Journal* (forthcoming), <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3388639>.

130 Giulia Schneider, ‘Disentangling Health Data Networks: A Critical Analysis of Articles 9(2) and 89 GDPR’ (17 September 2019) *International Data Privacy Law* <<https://academic.oup.com/idpl/advance-article/doi/10.1093/idpl/izp015/5571043?searchresult=1>>.

131 Wachter and Mittelstadt (n 96) 592.

132 In these regards, some clarifications have been provided by the Art. 29 Data Protection Working Party that has identified some examples in which companies carry out processing activities over personal data, without finalising them to individual decisions regarding natural persons, as in the case a business may wish to “classify its customers according to their age or gender for statistical purposes and to acquire an aggregated overview of its clients without making any predictions or drawing any conclusions about an individual. In this case the purpose is not assessing individual characteristics and is therefore not profiling”. So Art. 29 Data Protection Working Party, ‘Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679’ (3 October 2017, last modified 6 February 2018) <https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053>, 7.

133 Zarsky (n 105) 1008. It must however be said that in the context of big data analytics it is extremely difficult to identify secondary uses. So, Philipp Richter, ‘Big Data, Statistik Und Die Datenschutz-Grundverordnung’ (2016) 40 *Datenschutz und Datensicherheit*, 581, 585, highlighting the difficulties of detecting in which way the statistical models are employed, i.e. for which purposes and by which controllers.

134 Raum (n 117) 41. In this regard, a controller would need to have a different legal basis, such as consent or a task in the public interest, in order to employ a statistical model designed under the statistical research exemption. Stressing this point also, Wachter and Mittelstadt (n 96) 592 ff.

F. Conclusions: Research as an Efficiency Defence for Health Data Pools?

86 The above-traced framework leads to deeper considerations regarding the nature of the research exemption regarding the processing of health data under articles 9(2) lett. j); 5(1) lett. b); 6(4) and 89 GDPR within the system of the General Data Protection Regulation.

87 First of all, the detachment from the consent/control rule and the direct or possible (based on national legislation) derogation from some of data protection law's principles and data subjects' rights, suggests that the considered research exemption substantiates a regulatory paradigm that is not directly aligned to the General Data Protection Regulation's primary objective of the protection of data subjects' fundamental rights.

88 With regards to health data, this last objective is clearly satisfied by the prohibition of processing special categories of data under art. 9(1) para GDPR. As has been illustrated, however, this prohibition results to be largely weakened by the legitimate basis under art. 9(2) lett. j) GDPR that overall comes to liberalize the processing of special categories of personal data, as health data, for the purpose of scientific research.

89 This legal basis for the processing of special categories of personal data is characterized by a high degree of intrinsic and extrinsic vagueness¹³⁵: the intrinsic vagueness stems from the difficulties of clearly defining the notion of scientific and statistical research; the extrinsic vagueness is given by the Regulation's deferral of the definition of the conditions of processing for research purposes to Member States' legislation¹³⁶. Under these premises, art. 9(2) lett. j) GDPR appears to ultimately embed a substantially different rationale in respect to the other legal bases for the processing of special categories of data under art. 9(2) GDPR.

90 Indeed, the explicit consent under art. 9(2) lett. a) GDPR as a ground for processing is strictly rooted in data subjects' control and self-determination interests. This legal basis thus allows data subjects to autonomously and freely decide over their most

health information, in accordance to the individual fundamental rights of autonomy and dignity.

91 Under the public interest-related ground for processing under art. 9(2) lett. g) and i) GDPR, the processing of special categories of personal data is allowed for the achievement of higher societal and collective interests. The processing of special categories of data is in this case justified by higher interests, transcending individual data subjects' autonomy and self-determination expectations.

92 Conversely, the regulatory rationale of research as a basis for the processing of special categories of data, seems quite different. The research exemption under arts. 9(2) lett. j); 5(1) lett. b); 6(4) and 89 GDPR appears indeed to be the direct expression of what has been identified above as the second, internal market-oriented, pillar of the General Data Protection Regulation. Exactly in light of the General Data Protection Regulation's objective of promoting the free-flow of information within the internal market, the lawfulness of the processing of health data for research purposes under the mentioned provision can be read as a "safe harbor" for entities processing special categories of data, with the aim of stimulating innovation in data-driven markets, such as health data-driven markets¹³⁷.

93 In the practice, this means that the research exemption could work as a sort of efficiency defense under data protection law for the transfer and the processing of health data for research purposes, with subsequent market outcomes. Within the regulatory architecture of the General Data Protection Regulation, the research exemption thus seems to serve the original data protection law's internal market objectives.

94 Significant suggestions in this sense are given by recital 157 GDPR, which highlights the very functional nature of research, which is as an essential precondition for the "formulation and implementation of knowledge-based policy", and improves "the quality of life for a number of people" as well as "the efficiency of social services"¹³⁸. This holds especially true with respect to research over health data, whose great scientific value render

135 In this regard, Christiane Wendehorst, 'Of Elephants in the Room and Paper Tigers: How to Reconcile Data Protection and the Data Economy' in: Sebastian Lohsse, Reiner Schulze and Dirk Staudenmayer, *Trading Data in the Digital Economy: Legal Concepts and Tools* (Nomos/Hart Publishing, 2017) 327 ff.

136 Stressing this point also Zarsky (n 105) 1009.

137 Stressing a similar point in respect to the nature of the right to data portability, Inge Graef, Martin Husovec and Nadia Purtova, 'Data Portability and Data Control: Lessons from an Emerging Concept in EU Law' (2018) 19, 6 German Law Journal, 1359 ff. and also Graef, Gellert and Husovec (n 84) 16, highlighting that "data portability of Art. 20 GDPR is an example of an innovation policy embedded in data protection law". With regards to the research exemption, see Wachter and Mittelstadt (n 6) 592 ff.

138 Emphasis added.

them extremely important for the design of new products and services in the healthcare sector. In this perspective, the analyzed provisions regarding the processing of special categories of data for scientific and statistical research purposes are to be systemically aligned with other General Data Protection Regulation's provisions that appear to serve similar objectives.

- 95 In these regards, a parallelism emerges between the examined research exemption regarding health data and the right to data portability under art. 20 GDPR. This right has been indeed expressly welcomed by the Commission as a new means of promotion of the data economy, providing the data subject with the right to transfer his/her data from a service provider to another¹³⁹. Through this new right, thus, the data subject acquires an enhanced control over the data shared with businesses¹⁴⁰. Together with control rationales, however, the right to data portability ultimately stimulates data mobility across platforms, through data subjects' impulses¹⁴¹. From this perspective, hence, the right to data portability has been recently recognized by a strand of the literature as a tool for data-innovation and the promotion of the free-flow of personal-information¹⁴². However, the right to data portability is still based on data subjects' control over their data in respect to processing platforms, since the flow of data is enacted only upon the data subjects' determinations. To the very contrary, under the research exemption for the processing of special categories of data under arts. 9(2) lett. j); 5(1)

lett. b); 6(4) and 89 GDPR data subjects appear to be significantly excluded from the control over their processed health data.

- 96 Under these provisions, by establishing a special regime regarding processing activities over health data carried out for research purposes, the General Data Protection Regulation provides normative grounds for incentivising data-driven research activities, in consistency with the European Commission's promotion of digital health and the free-flow of information within the internal market.
- 97 Hence, the General Data Protection Regulation appears to reflect aspects of economic regulation, which ultimately facilitate the creation of a market of personal health data and in this way set the conditions for the efficient functioning of other markets¹⁴³, such as the one for digital medical devices and pharmaceuticals.
- 98 From a regulatory standpoint, thus, the General Data Protection Regulation's research exemption regarding the processing of special categories of personal data appears to be not a *data protection* rule but rather a *rule of the data economy*, which nonetheless addresses data protection concerns, expressed in the requirement of the enactment of safeguards for the respect of data subjects' fundamental rights¹⁴⁴. This acknowledgement leaves open the question whether the safeguards required under arts. 9(2) lett. j) and 89 GDPR for the protection of data subjects' fundamental rights in the context of health data pools and the related research activities are sufficient; or whether there is the need to integrate these with other regulatory safeguards, provided for example by competition law or ethical guidelines.

139 European Commission, 'Commission Staff Working Paper Impact Assessment Accompanying the Document Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of such Data (General Data Protection Regulation) and the Directive of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data by Competent Authorities for the Purposes of Prevention, Investigation, Detection or Prosecution of Criminal Offences or the Execution of Criminal Penalties, and the Free Movement of such Data, SEC (2012) 72/2' (2012) SEC(2012) 72/2 <<https://ec.europa.eu/transparency/regdoc/rep/2/2012/EN/SEC-2012-72-2-EN-MAIN-PART-1.PDF>>, 53.

140 Josef Drexler, 'Data Access and Data Control in the Era of Connected Devices, Study on Behalf of the European Consumer Organisation BEUC' (27 April 2018) Beuc <https://www.beuc.eu/publications/beuc-x-2018-121_data_access_and_control_in_the_area_of_connected_devices.pdf>, para 30.

141 Graef, Gellert and Husovec (n 84) 3.

142 Graef, Husovec and Purtova (n 137) 1396 ff.

143 This is highlighted from a general perspective by Lynskey (n 99) 76-77.

144 For a distinction between the rules regarding data protection and data economy, see Wendehorst (n 135) 332.