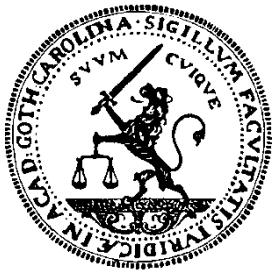

FACULTY OF LAW

LUND UNIVERSITY

Anna Bockasten

Health Data in Data-driven Research:
Processing under the GDPR's Scientific Research Exemption

JURM02 Graduate thesis

Graduate thesis, Master of Laws program

30 higher education credits

Supervisor: Valentin Jeutner

Semester: Spring 2024

Table of contents

Summary

The digitalization of the healthcare sector has resulted in an increasing source of health data, enabling the implementation of artificial intelligence (AI) in healthcare. There is great optimism that AI will have a significant impact on all areas of healthcare. The processing of health data is generally prohibited by the General Data Protection Regulation (GDPR). However, article 9(2)(j) GDPR provides for an exemption when the processing is carried out for scientific research purposes. The scientific research regime in the GDPR further includes exceptions from principles and obligations and allows for derogations from several data subjects' rights. The scope of the scientific research exemption is not entirely clear, as the GDPR does not contain a binding definition of "scientific research purposes" and as rules may vary by Member State.

The thesis examines the legal impact of the GDPR in relation to Swedish companies engaging in data-driven research, by asking to what extent Swedish companies can claim the scientific research exemption in Article 9(2)(j) GDPR when processing health data. In Swedish law, ethical review pursuant to the Swedish Ethical Review Act is required to process health data under the scientific research exemption. Therefore, the relationship between "scientific research purposes" in the GDPR and "research" as defined in the Ethical Review Act is examined. The thesis concludes that neither the GDPR nor the Ethical Review Act preclude private entities or activities that are undertaken with a commercial interest. As the definition of research in the Ethical Review Act focuses on the acquirement of new knowledge and the theoretical and/or practical value of research, the thesis argues that the definition sets forth a higher threshold of what constitutes research than the GDPR. Companies that are primarily driven by commercial interests might have difficulties clarifying the scientific value of their activities. A disadvantage is that the definition of research in the Ethical Review Act and its territorial scope do not align with that of the GDPR, creating a fragmented legal framework within the EU.

Secondly, the thesis asks how the scientific research regime and its implementation in Swedish law balance the interests of data subjects against the interests of controllers, and how this balance might affect data-driven research. It concludes that the scientific research regime appears at first sight to shift the balance of interests significantly in favor of the controller. However, it is often required that the provisions' application would render impossible or seriously impair the achievement of scientific research, thereby narrowing the scope of the framework. The thesis highlights Sweden's passive stance in terms of legislation, for example by refraining from introducing the possibility to derogate from certain rights of the data subject. While this may adversely affect the flexibility of companies engaging in data-driven research, the biggest challenge is to overcome the conflict between the GDPR and research involving substantial amounts of personal data.

Sammanfattning

Digitaliseringen av hälso- och sjukvårdssektorn har resulterat i en ständigt växande tillgång till hälsodata, vilket möjliggör implementeringen av artificiell intelligens (AI) inom hälsosektorn. Det finns en stor tilltro till att AI kommer att ha en betydande inverkan på alla delar av hälso- och sjukvården. Behandlingen av hälsodata är i allmänhet förbjuden enligt EU:s dataskyddsförordning (GDPR). Artikel 9.2 j GDPR föreskriver dock ett undantag när behandlingen utförs för vetenskapliga forskningsändamål. Bestämmelserna om vetenskaplig forskning i GDPR medger vidare undantag från principer och skyldigheter samt tillåter undantag från vissa av den registrerades rättigheter. Omfattningen av undantaget för vetenskaplig forskning är däremot inte klarlagt, eftersom GDPR saknar en bindande definition av ”vetenskapliga forskningsändamål” och eftersom reglerna kan variera mellan medlemsstaterna.

Uppsatsen undersöker GDPRs rättsliga inverkan på företag som bedriver dataintensiv forskningsverksamhet, genom att utreda i vilken utsträckning svenska företag kan åberopa undantaget för vetenskaplig forskning i artikel 9.2 j GDPR vid behandling av hälsodata. Enligt svensk rätt föreskrivs etikprövning enligt etikprövningslagen som ett krav för att behandla hälsodata med stöd av undantaget. Förhållandet mellan ”vetenskapliga forskningsändamål” i GDPR och definitionen av ”forskning” i etikprövningslagen utreds därför. Uppsatsen drar slutsatsen att varken GDPR eller etikprövningslagen utesluter privata aktörer eller kommersiella intressen. Eftersom definitionen av forskning i etikprövningslagen fokuserar på inhämtandet av ny kunskap och forskningens teoretiska och/eller praktiska värde, anses den uppställa högre krav för vad som utgör forskning än GDPR. Företag som primärt drivs av kommersiella intressen kan ha svårt att tydliggöra det vetenskapliga värdet i sin verksamhet. En nackdel är att definitionen av forskning i etikprövningslagen och dess territoriella tillämpningsområde inte överensstämmer med GDPRs, vilket skapar ett fragmenterat regelverk inom EU.

Uppsatsen undersöker vidare hur GDPRs regelverk kring vetenskaplig forskning och implementeringen av detta i svensk rätt balanserar den registrerades intressen mot den personuppgiftsansvariges, samt hur denna avvägning kan påverka dataintensiv forskningsverksamhet. Uppsatsen drar slutsatsen att bestämmelserna vid första anblick tydligt är till förmån för den personuppgiftsansvarige. Det krävs dock ofta att tillämpningen av bestämmelserna skulle göra det omöjligt eller mycket svårare att uppfylla de särskilda ändamålen, vilket begränsar tillämpningsområdet. Uppsatsen belyser att Sverige har agerat passivt i lagstiftningshänseende, till exempel genom att avstå från att införa möjligheten att föreskriva undantag från vissa av den registrerades rättigheter. Även om detta kan ha en negativ inverkan på flexibiliteten för företag som bedriver dataintensiv forskning är den största utmaningen att hantera den konflikt som finns mellan GDPR och dataintensiv forskningsverksamhet.

Preface

I would like to thank my supervisor Valentin Jeutner for his valuable insights that have been very helpful for the work with this thesis.

I also want to thank my mom, dad and Madeleine for your support throughout these five years.

A special thank you to my sister Chloé, my greatest support.

Lund, May 21, 2024
Anna Bockasten

Abbreviations

CJEU	Court of Justice of the European Union
ECtHR	European Court of Human Rights
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
EU	European Union
GDPR	General Data Protection Regulation
OECD	Organisation for Economic Co-operation and Development
TFEU	Treaty on the Functioning of the European Union
WP29	Article 29 Data Protection Working Party

1 Introduction

1.1 Background

The digitalization of the healthcare sector in the last decade has resulted in a continuously growing source of data.[1] Through medical research, electronic health records, the delivery of healthcare services, and the use of electronic devices and health applications, health data is constantly collected, either intentionally or unintentionally. The use of search engines such as Google to assess symptoms or concerns generates large volumes of data, which is of great value to companies wanting to use it themselves or sell it to third parties.[2] While collected data offers useful information for research, the sharing and use of people's health data can lead to controversial scenarios. One example is the Google DeepMind patient data deal, which was later found to fail to comply with data protection law.[3] In 2015, Google DeepMind struck a deal with the Royal Free Hospital under the National Health Service (NHS), which granted DeepMind access to health data from over 1.6 million patients in order to develop a medical device application for kidney disease.[4] Other companies have made it their business plan to resell data. One example is 23andMe, a company offering genetic testing kits, while asking people for their consent to offer their genetic and health data for research. The company has entered into several deals with large pharmaceutical and biotech companies, for example, a million-dollar deal with the biotech company Genentech that allows it to use its database.[5]

The current access to large health data sets enables the implementation of artificial intelligence (AI) in healthcare.[6] There is great optimism that AI will have a significant impact on all areas of healthcare, from clinical applications in areas such as imaging and diagnostics, to the use of health apps to assess an individual's symptoms. Already, there is evidence that AI algorithms perform as well or better than humans in several fields, for example by analyzing medical images or providing a prognosis of the disease process, based on symptoms and biomarkers in electronic medical records.[7] AI will play an

[1] Bohr and Memarzadeh (2020), p. 10.

[2] Ibid., p. 12-13.

[3] Information Commissioner's Office (ICO) (2017), p. 1.

[4] Meszaros et al. (2020), p. 93. DeepMind is a British AI company that was bought by Google in 2014.

[5] Bentzen (2020), p. 343; Herper, M. "Surprise! With $60 Million Genentech Deal, 23andMe Has a Business Plan," Forbes (Jan 6, 2015).

[6] Bohr and Memarzadeh (2020), p. 16, 26. Artificial intelligence, or AI, refers to technology that enables computers and machines to display human intelligence.

[7] Ibid., p. 25; Gerke et al. (2020), p. 295.

important part in assisting people to maintain good health and will result in earlier diagnosis, tailored treatments, and more efficient monitoring.[8]

An example of a Swedish company operating in the fields of medicine, data science and software development is AI Medical Technology, which has developed a diagnostic decision support system that classifies skin cancer using image analysis combined with deep learning.[9] The product is powered by AI, which has been developed and trained on dermatoscopic images of patients' skin lesions together with patient data.[10] Another Swedish company that uses AI to improve people's health is Deversify, a research and development company that develops and commercializes mobile electronic devices and apps to monitor biomarkers for individualized health. Furthermore, the company collects, analyzes, and conducts research on the data generated by the products, with new knowledge being communicated to both consumers and the academic research community.[11]

While AI has the potential to transform the healthcare sector, the use of health data in data-driven research raises concerns about data protection and privacy, given the sensitive nature of these data, and about ensuring that data is used for lawful activities.[12] Health data has traditionally been considered sensitive, based on a general perception that it holds some of the most private and intimate parts of ourselves and if revealed, it might lead to stigmatization, for example in the case of mental health conditions such as schizophrenia.[13] Unlike leaked credit card details or a hacked online account, the harm a person may experience if details about their health conditions are released cannot be hindered by blocking access to a bank account or changing a password.[14] The mere knowledge that information about one's health is "out-there," at risk of being exploited by others, can be enough to cause emotional distress.[15] Another concern is the risk of discrimination. For example, insurance companies may raise their insurance costs for health, life, or disability insurance or not accept individuals at all based on collected health data.[16] While concerns in the form of insurance loss or emotional distress can also occur with smaller collections of data, big data tends to increase the number of people affected.[17]

[8] Bohr and Memarzadeh (2020), p. 26.

[9] Deep learning is a subset of machine learning (which falls under AI) that relies on algorithms that can generate patterns in data once they are fed with enough training examples.

[10] Dermalyser, "About" and "Dermalyser", https://www.aimedtech.com.

[11] Deversify, "About us" and "R&D", https://deversify.com.

[12] Hlávka (2020), p. 235.

[13] Bygrave and Tosoni, "Article 4(15). Data Concerning Health" (2020), p. 218-219; Determann (2020), p. 257.

[14] Determann (2020), p. 257.

[15] Nicholson Price 2nd and Glenn Cohen (2019), p. 38.

[16] Determann (2020), p. 258. For example, the ECtHR case *Vukota-Bojić v. Switzerland* concerned secret surveillance by an insurance company of an applicant in order to question her level of disability.

[17] Nicholson Price 2nd and Glenn Cohen (2019), p. 38. Big data is characterized "as any collection of data that is so large in terms of both volume and complexity" (see Bohr and Memarzadeh [2020], p. 12).

Furthermore, widely disclosure of health data might affect the trust in health care providers and deter people from voluntarily participating in research.[18]

The view of health data as sensitive is also reflected in longstanding rules on patient confidentiality concerning medical records kept by doctors. Furthermore, health data has consistently been given special status in EU laws on data protection by including it as "sensitive data." This tradition has continued in the General Data Protection Regulation (GDPR).[19] The processing of health data is generally prohibited by the GDPR, however, there are exemptions, such as if the data subject has given its explicit consent to the processing or if it is done for scientific research purposes: the scientific research exemption.[20]

Processing for scientific research purposes is of particular interest not only because it provides an exemption to the prohibition of the processing of sensitive data but also because it allows for exceptions from certain principles and obligations, and allows for derogations from the data subject's right to access, rectification, restriction, and objection. Thus, scientific research is privileged in the GDPR and provides for a favorable regime, making it attractive not only to academic researchers but also to commercial entities engaging in for-profit research.[21]

The GDPR does not contain a binding definition of what is considered "scientific research purposes," but should, according to Recital 159 GDPR, be interpreted broadly and include for example technological development and studies conducted in the public interest in the area of public health. Furthermore, the processing must be based on either Union or Member State law, which means that different rules may apply depending on the Member State in question. This raises questions not only about the scope of the scientific research exemption in the GDPR and which entities can claim it, but also about the balance between data protection and advances and innovations in the healthcare sector. What activities can benefit from the scientific research exemption, and how do commercial interests affect this assessment?

1.2 Purpose and Research Questions

The thesis aims to examine the legal impact of the GDPR in relation to Swedish companies engaging in data-driven research, by focusing on the "scientific research exemption" in the GDPR. As different rules may apply depending on the Member State in question, the thesis will consider this relationship

[18] Bygrave and Tosoni, "Article 4(15). Data Concerning Health" (2020), p. 218.
[19] Ibid., p. 218-219.
[20] Article 9 GDPR.
[21] Bentzen (2020), p. 342; Ducato (2020), p. 4; Kruus (2023), p. 65.

between the GDPR and Member State law by using Swedish law as a practical example.

The thesis aims to answer the following questions:

- To what extent can Swedish companies claim the scientific research exemption in Article 9(2)(j) GDPR when processing health data?

- How do the GDPR's scientific research regime and its implementation in Swedish law balance the interests of data subjects against the interests of controllers, and how might this balance affect data-driven research?

1.3 Method and Material

To answer the two research questions, the thesis uses a traditional legal method. The sources of law that form the basis of the analysis are the GDPR, which as a regulation is directly applicable in Member States, and the Swedish Act Concerning Ethical Review of Research Involving Humans (Lag [2003:460] om etikprövning av forskning som avser människor) (the Ethical Review Act), which applies to research involving the processing of special categories of personal data, as defined in the GDPR. The thesis will also examine the preparatory work for the Ethical Review Act and the preparatory work in relation to the legislative amendments that were made as a result of the entry into force of the GDPR. Furthermore, guidelines by the Swedish Ethical Review Authority, which examines applications for ethical review for research conducted in Sweden, will be considered, as they can assist in shaping the understanding of how the Ethical Review Act is applied in practice.

The thesis will consider the relationship between EU law and domestic law. Guidelines by the European Data Protection Board (EDPB) and the Article 29 Working Party (WP29) will be included. The EDPB is an independent European body that adopts guidelines to ensure a consistent application of the GDPR. It replaced the WP29, which was established by Directive 95/46/EC (GDPRs predecessor) and dealt with issues relating to the protection of privacy and personal data. Guidelines issued by the WP29 may still carry relevance in relation to provisions in the GDPR that resemble those in the Directive. Some of the guidelines by the WP29 have also later been endorsed by the EDPB. The guidelines by the EDPB and WP29 as soft law are not binding, and it is therefore possible to question their authority. Soft law has traditionally not carried too much weight but there has recently been a shift as the Court of Justice of the European Union (CJEU) has explicitly referred to them when interpreting legal provisions.[22] In practice, despite being non-

[22] In C-322/88 *Grimaldi*, para. 3, the CJEU stated that "since recommendations cannot be regarded as having no legal effect at all, the national courts are bound to take them into consideration in order to decide disputes submitted to them."

binding, guidelines such as the ones issued by the EDPB and WP29 can often carry significant normative value.[23]

The GDPR leaves a lot of room for interpretation in relation to processing for scientific research, as the definition of scientific research purposes is included in the non-binding part of the regulation and as the recital calls for a broad interpretation. Furthermore, neither the CJEU nor the EDPB has issued statements on the breadth of the scientific research exemption. The thesis will therefore draw from the discussion by legal scholars regarding the scientific research exemption, which will form the basis of the analysis concerning the relationship between the GDPR and Swedish law.

1.4 Current Research

Several scholars have discussed the breadth of the scientific research exemption in the GDPR. Pormeister has written about the research exemption in regard to genetic data, stating that the broad interpretation that the GDPR calls for could mean that private business-to-consumer companies, who hold large amounts of data, may use it for research purposes.[24] In a 2018 paper, Mészáros and Ho write that it is unclear how the scientific research exemption is applied in the "corporate context," for example regarding product improvement and data analytics. In 2021, the same authors stated that the GDPR's broad definition of science allows for private companies to conduct commercial research.[25] They note, however, that it can be challenging to identify the research component in software development, and that an upgrade or change of an already existing program or system may only be classified as research if it results "in an increase in the stock of knowledge."[26]

The focus among several of the scholars mentioned has been whether or not private companies should benefit from the research exemption. Pormeister argues that the exemption shifts the balance of interests rather heavily in favor of the data processor/controller, which leaves the data subject with limited control over the processing of their genetic data for scientific research.[27] Mészáros and Ho argue that the lack of a binding legal definition of scientific research in the GDPR may result in different interpretations and forum shopping, which can undermine the privacy of individuals, as Member States want to be at the forefront of scientific research.[28] In another paper, Mészáros and Ho argue that commercial AI research should not benefit from the research exemption in the GDPR without a public interest and similar safeguards as

[23] Riesenhuber (2021), p. 249 ff.
[24] Pormeister (2017), p. 145.
[25] Mészáros and Ho (2021), p. 1.
[26] Ibid., p. 8.
[27] Pormeister (2017), p. 146.
[28] Mészáros and Ho (2018), p. 407.

academic research, as commercial research contains a lower ethical standard and because "corporate secrecy" creates barriers for oversight.[29]

These arguments can be said to be part of a larger discussion about the GDPR's impact on research and innovation. For example, Determann argues that the discretion given to EU Member States to legislate derogations from the GDPR enables the creation of a legal patchwork that makes it more difficult for research institutions and companies to undertake international studies or exchange data across borders.[30] Comandè and Schneider instead argue that European data protection law does not hinder but rather encourages data-driven research.[31]

The implementation of the research exemption has been discussed regarding for example Estonia, England, and Germany.[32] While Slokenberga et. al. provide a brief overview of the implementation in Sweden, the analysis is centered around the framework for the European Health Data Space.[33] Holtz and Ledendal have researched commercial processing of data, by looking at the overlap of the GDPR and rules governing marketing.[34] The extent of the scientific research exemption in Swedish law has, however, not been researched further.

1.5 Delimitations

The thesis will focus on the perspective of companies and their ability to pursue data-driven research involving health data. While the thesis discusses the balance between the interests of data subjects and the interests of controllers, it is mostly interested in the ways in which the scientific research exemption can be used by companies to justify activities that would otherwise conflict with the interests of data subjects. The perspective of data subjects will therefore not be given the same attention as the perspective of companies.

Article 9 GDPR also contains exemptions for the processing of sensitive data for health care purposes and the processing of sensitive data for reasons of public interest in the area of public health.[35] As in regard to the processing for health care purposes under Article 9(2)(h), the provision does not cover processing for purposes of medical research and the exemption will therefore be left out of the thesis.[36] As for the exemption concerning public interest in the area of public health, it is understood to be a narrow exemption intended for

[29] Mészáros and Ho (2021), p. 9-10.
[30] Determann (2020), p. 241.
[31] Comandè and Schneider (2022), p. 4.
[32] See Pormeister (2017) and Mészáros and Ho (2018).
[33] Molnár-Gábor et al. (2022), p. 271.
[34] Holtz and Ledendal (2020).
[35] Article 9(2)(h) and (i) GDPR.
[36] Georgieva and Kuner (2020), p. 379.

use by public health authorities and non-governmental organizations.[37] As this thesis examines the legal impact of the scientific research exemption in relation to companies engaging in data-driven research, it will also not be explored further.

Sweden will be used as a practical example of how the scientific research exemption has been regulated in Member State law. However, the aim is not to provide a comprehensive account of Sweden's data protection legislation. The same applies to the GDPR. As it is a substantial piece of legislation, the thesis will only cover the provisions that are relevant in regard to the purpose of the research, with a particular focus on Article 9(2)(j) GDPR.

1.6 Outline

The thesis consists of seven chapters. The analysis will follow directly in the concluding section of each chapter. While the concluding section of Chapter 2 mainly summarizes the material, the concluding sections of the subsequent chapters will aim to answer the two research questions.

Chapter 2 will provide the reader with an introduction to the GDPR and highlight some key provisions and definitions of the legislation.

Chapter 3 will focus on the scientific research exemption in Article 9(2)(j) GDPR and its applicability. It will further outline the debate among scholars on the extent of the scientific research exemption in relation to data processing by companies.

Chapter 4 will draw from the conclusions made in Chapter 3 and examine the relationship between "scientific research purposes" in the GDPR and "research" as defined in the Swedish Ethical Review Act, to discuss the effect it may have on the ability of companies to claim the scientific research exemption when processing health data.

Chapter 5 will outline the legal grounds on which the processing of health data must be based, apart from being covered by an exception in Article 9(2) GDPR, as this affects the ability of companies to process health data.

Chapter 6 will mainly focus on the second research question, namely the balance between the interests of data subjects and controllers. It will further outline the regulatory framework around processing for scientific research purposes.

[37] Ibid., p. 380.

Chapter 7 will bring together the conclusions that have been made in regard to the research questions and offer some reflections in relation to the findings of the thesis.

2 An Introduction to the GDPR: Key Provisions and Definitions

The General Data Protection Regulation (GDPR) entered into force in May 2018, replacing the 1995 Data Protection Directive.[38] It marks the first thorough reform of the legal framework on data protection within the EU since the adoption of the Directive. The GDPR recognizes data protection of individuals as a fundamental right, which is also enshrined in several EU instruments,[39] including the Charter of Fundamental Rights of the European Union.[40] The aim of the GDPR is to uphold the right to data protection, while also ensuring the free movement of personal data within the EU, in order to facilitate cooperation in areas such as economy and science.[41]

This chapter will provide the reader with an introduction to the GDPR to better the understanding of what situations entail the application of the GDPR. It will also address principles that apply to the processing of personal data and the definition of "data concerning health" in the GDPR.

2.1 Material Scope

Article 2 of the GDPR contains the material scope of the regulation: it applies to the processing of personal data, which encompasses any information that is related to an identified or identifiable natural person.[42] The CJEU has stated that personal data does not only cover data that is considered sensitive or private but can potentially include all kinds of information as long as it relates to a particular person.[43] Over the years, the CJEU has found a diverse range of types of information to constitute personal data, such as a person's telephone number or information on their working conditions and hobbies,[44] images of persons recorded on video camera,[45] and IP addresses in certain situations.[46] Bygrave and Tosoni point out that the intentionally broad and flexible definition of "personal data" in the GDPR may have its drawbacks. Since the law is given an almost enormous scope, it may affect its practical application in terms of compliance and enforcement. This cost becomes more

[38] 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.
[39] See for example Article 16 of the Treaty on the Functioning of the European Union.
[40] Article 8 of the EU Charter of Fundamental Rights.
[41] Article 1 GDPR.
[42] Article 2(1) GDPR; Article 4(1) GDPR.
[43] Case C-434/16, *Nowak*, para. 34-35.
[44] Case C-101/01, *Lindqvist*, para. 27.
[45] Case C-212/13, *Ryneš*, para. 22; Case C-345/17, *Buivids*, para. 32.
[46] Case C-582/14, *Breyer*, para. 49.

significant in the era of "Big Data Analytics," when data that has previously appeared to be anonymous no longer is.[47]

The GDPR does not apply to the processing of anonymous data or data that has been anonymized in a manner that ensures that the data subject is no longer identifiable.[48] For example, data can be anonymized by removing personal identifiers such as names and social security numbers and by modifying other types of "identifiers," such as deleting the names of relatives.[49] To determine whether a person is identifiable, the GDPR states that consideration should be given to all means that are reasonably likely to be used, including considering factors such as cost and the amount of time that the identification would require.[50] The CJEU has expressed that a person cannot be considered identifiable if the identification of the data subject is prohibited by law or practically impossible because it would require a disproportionate effort in terms of time, cost, and manpower, thereby rendering the risk practically insignificant.[51] However, data can be personal even if the controller cannot correlate the data to a particular person without help from other sources.[52] In *Breyer*, the CJEU stated that "it is not required that all the information enabling the identification of the data subject must be in the hands of one person."[53]

For the GDPR to be applicable, personal data must be processed. "Processing," like personal data, is defined broadly and includes, for example, collection, recording, organization, structuring, and storage of personal data.[54] The WP29 has stated that data collection without recording, or storage still entails the application of the data protection legislation.[55] Data processing would therefore appear to be covered by the definition in the GDPR, irrespective of its duration, the amount of data processed, and the actual recording of personal data.[56]

[47] Bygrave and Tosoni, "Article 4(1). Personal Data" (2020), p. 113. The authors do not develop their reasoning about the interaction between big data and anonymization. One risk however with the introduction of AI technologies is that AI is likely to increase the ability to reidentify individuals in anonymized datasets (see Hlávka, 2020), which may be what the authors are referring to. This was shown, for example, in a study in which over 90% of adult individuals were successfully reidentified using AI to match data collected from wearable devices to individuals (see Liangyuan et al., 2018).

[48] Recital 26 GDPR.

[49] Ohm (2010), p. 1703.

[50] Recital 26 GDPR.

[51] C-582/14, *Breyer*, para. 46.

[52] Bygrave and Tosoni, "Article 4(1). Personal Data" (2020), p. 110-111.

[53] C-582/14, *Breyer*, para. 43.

[54] Article 4(2) GDPR.

[55] WP29, "Opinion 1/2015 on Privacy and Data Protection Issues Relating to the Utilisation of Drones" (2015), p. 7, fn. 13.

[56] Bygrave and Tosoni, "Article 4(2). Processing" (2020), p. 119.

2.2 Territorial Scope

Article 3 covers the territorial scope of the regulation and contains three situations that trigger the application of the GDPR. To understand when the obligations of the GDPR are invoked, the definitions of "controller" and "processor" are essential. A controller is a natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data, while a processor processes personal data *on behalf of* the controller.[57] Data controllers and processors are obligated to protect personal data under the GDPR regardless of if they are public or private entities.[58]

The definition of "controller" must be understood in light of the legislator's aim of placing the main responsibility for data protection on the entity that has the actual control over the data processing, which entails considering not only legal formalities but also factual circumstances.[59] One of the requirements to qualify as a controller under Article 4(7) GDPR is that the entity determines the purposes and means – the "why" and "how" – of the processing activities. The WP29 has stated that an entity that decides the purposes of the processing qualifies as controller, while an entity that decides the means of the processing only qualifies as controller as long as it decides on "essential elements of the means," such as which data shall be processed or for how long. Technological and organization questions, such as for example which hardware or software that is to be used for the processing, can be delegated to processors and does not trigger the qualification as controller, in the view of the WP29.[60]

The role of the "processor" is closely connected to that of the "controller," as it arises from a delegation or "outsourcing" of tasks decided by the controller. A processor must be legally separate from the controller. The relationship between the two entities is therefore to be distinguished from an employer-employee relationship, as an employee that processes personal data to fulfill their obligations towards the employer should not be considered a processor.[61] The WP29 has also highlighted that the role of the processor does not derive "from the nature of an entity processing data but from its concrete activities in a specific context."[62] The processor must comply with the controller's instructions regarding the purposes and means of the processing. An entity is therefore only a processor insofar as it acts within the scope of responsibility established by the controller.[63] When a processor acts outside this

[57] Article 4(7) and (8) GDPR.
[58] Krzysztofek (2021), p. 37.
[59] Bygrave and Tosoni, "Article 4(7) Controller" (2020), 148. This was also noted by the WP29, see WP29 (2010), p. 9.
[60] WP29 (2010), p. 14.
[61] Bygrave and Tosoni, "Article 4(8). Processor" (2020), p. 159.
[62] WP29 (2010), p. 25.
[63] Bygrave and Tosoni, "Article 4(8). Processor" (2020), p. 160.

scope and thus starts to determine the purposes and means of the processing, it ceases to be a processor and assumes the role of controller, as follows from Article 28(10) GDPR.

For example, a company looking to enter the medtech industry wants to develop an AI-based tool for breast cancer screening. In order to train the model, the company hires a software developer and provides it with clear instructions on, for example, what data that should be collected, for how long it should be processed, and who should have access to the data. In this case, the company, as it decides why and how data should be processed, is the controller, and the software developer is the processor, as it processes data on behalf of the company in accordance with the company's instructions.

As described above, Article 3 contains three situations concerning the processing of personal data by a controller or processor that trigger the application of the GDPR. Firstly, the GDPR applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the EU, regardless of whether the processing takes place in the EU or not.[64] As highlighted by Recital 14, the application of the GDPR is not dependent on a person's nationality or place of residence. The EDPB has emphasized that "it is the presence, through an establishment, of a data controller or a processor in the EU and the fact that a processing takes place in the context of the activities of this establishment" that leads to the application of the GDPR to the processing activities.[65] This was clarified by the EDPB with the following example:

> A pharmaceutical company with headquarters in Stockholm has located all its personal data processing activities with regards to its clinical trial data in its branch based in Singapore. In this case, while the processing activities are taking place in Singapore, that processing is carried out in the context of the activities of the pharmaceutical company in Stockholm i.e. of a data controller established in the Union. The provisions of the GDPR therefore apply to such processing, as per Article 3(1).[66]

Secondly, the GDPR applies to a controller or processor that is not established in the EU. For the GDPR to be applicable in these situations, the processing must be of personal data of data subjects who are in the EU, and the processing activities must be related to the offering of goods or services to such data subjects or to the monitoring of their behavior, as far as their behavior takes place within the EU.[67] Thirdly, the GDPR applies to a controller not

[64] Article 3(1) GDPR.

[65] EDPB, "Guidelines 3/2018 on the Territorial Scope of the GDPR (Article 3)" (2019), p. 9.

[66] Ibid., p. 9-10.

[67] Article 3(2) GDPR.

established in the EU, but in a place where Member State law applies by virtue of public international law.[68]

2.3 Principles of Data Processing

Article 5 of the GDPR contains principles for the processing of personal data. These principles form the basis for data protection law and all EU laws that regulate the protection of personal data must adhere to these principles. Furthermore, they act as guidelines for the interpretation of other legal provisions.[69] In accordance with the principles, personal data must be:

- Processed lawfully, fairly, and in a transparent manner in relation to the data subject (lawfulness, fairness, and transparency);
- Collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with the initial purposes (purpose limitation);
- Adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed (data minimization);
- Accurate and where necessary, kept up to date (accuracy);
- Kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed (storage limitation);
- Processed in a manner that ensures appropriate security of the personal data (integrity and confidentiality).

The controller is responsible for and must be able to demonstrate compliance with the principles (accountability).[70]

2.4 Special Categories of Personal Data

Certain categories of personal data have acquired an additional layer of protection in the GDPR. These are data that according to Recital 51 GDPR are particularly sensitive by nature and require additional protection as their processing may pose significant risks to fundamental rights and freedoms. The special categories of personal data are listed in Article 9(1) GDPR and include data that reveals racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data for the

[68] Article 3(3) GDPR.
[69] Krzysztofek (2021), p. 59.
[70] Article 5(1)(a)-(f) and (2) GDPR.

purpose of uniquely identifying a person, data concerning health or data concerning a person's sex life or sexual orientation.[71]

As a general rule, the processing of special categories of personal data is prohibited under the GDPR, however, there are exemptions.[72] As explained above, processing for scientific research purposes is one exemption allowing the processing of special categories of data.[73] Another exemption is if the data subject has given explicit consent to the processing for one or more specified purposes.[74] Regarding genetic data, biometric data, or data concerning health, Member States may impose additional conditions, including limitations, to the processing of this type of data.[75]

2.4.1 Data Concerning Health

One category of personal data is data concerning health, which is defined in Article 4(15) GDPR as all personal data that relates to a person's physical or mental health, including data that relates to the provision of health care services, which reveals information about a person's health status. It follows from Recital 35 that data concerning health should include any data that reveals information about a person's past, present, or future physical or mental health status. This includes a number or symbol that has been assigned to a person to uniquely identify them for health purposes; information derived from testing or examination of a body part or bodily substance; and any information on, for example, a disease, disability, disease risk, medical history, clinical treatment or physiological or biomedical state of the data subject, irrespective of its source, for example from a physician or other health professional, a hospital, a medical device or an in vitro diagnostic test.

The WP29 has provided substantial guidance on what is covered by the concept of "health data" and has specifically addressed to what extent health data in apps and devices qualifies as such. The WP29 states that while medical data generated within a professional medical context, including data generated by devices or apps used in this context, is a clear example of health data, the term includes far more than such medical data. Health data may include information derived from devices that analyze a person's urine or blood, and apps that measure blood pressure or heart rate, regardless of whether the testing is performed by professionals or with devices and apps that are commercially available.[76]

[71] Article 9(1) GDPR.
[72] Article 9(1) and (2) GDPR.
[73] Article 9(2)(j) GDPR.
[74] Article 9(2)(a) GDPR.
[75] Article 9(4) GDPR.
[76] WP29, "Annex – Health Data in Apps and Devices" (2015), pages are not numbered.

However, the WP29 assumes that certain personal data generated by lifestyle apps and devices should generally not be considered data concerning health within the meaning of the GDPR. The WP29 provides an app that counts steps during a walk as an example of a collection of data, that without any additional information about the data subject, or a medical context in which the data is collected, does not constitute health data in the sense that it needs protection as a special category of data. However, raw personal data with a "relatively low privacy impact" can easily transform into health data when the data set can be used to determine an individual's health status.[77] In this regard the WP29 provides another example:

> A single registration of a person's weight, blood pressure or pulse/heart rate (if not excessive in absolute terms), at least without any further information about age or sex, does not allow for the inference of information about the actual or likely future health status of that person. However, that aspect measured over time, especially in combination with age and sex, may be used to determine a significant aspect of an individual's health, such as the health risks related to obesity or an illness causing a significant loss of weight, high/low blood pressure, arrhythmia, etc.[78]

Zarsky argues that big data analytics - which involves the processing and interpretation of data to extract valuable information - challenges the ability to distinguish between special and other categories of personal data, as an analysis involving the processing of "regular" personal data quickly can shift into one relying on sensitive data.[79] Wouters et al. explain this further by pointing to the fact that further expansion of social media, smart devices, and AI in people's everyday lives will expand the amount of potential health data. Through the interconnection and integration of data concerning dietary habits, sleep patterns, and social media, "regular" data (what the WP29 refers to as raw personal data with "relatively low privacy impact") can quickly turn into sensitive data requiring additional protection.[80] Zarsky argues that the need under the GDPR to differentiate between different types of personal data complicates big data analytics, as the shift from one category to another requires the application of different provisions. Furthermore, beyond complicating the process, big data processes may undermine the distinction. If data sets containing personal data can produce sensitive data, the distinction may appear "almost artificial."[81]

[77] Ibid.
[78] Ibid.
[79] Zarsky (2017), p. 1013.
[80] Wouters et al. (2021), p. 209.
[81] Zarsky (2017), p. 1013.

2.5 Summary and Concluding Remarks

This chapter introduced the reader to some key provisions and definitions in the GDPR. Firstly, it presented the material scope of the regulation, the processing of personal data, which is extensive, as the definitions of personal data and processing are to be interpreted broadly. As highlighted by the CJEU, personal data does not only cover sensitive or private data, but essentially all types of data as long as it relates to an identifiable person. The GDPR therefore does not apply to anonymous data or data that has been anonymized, for example, data where identifiers such as name and social security numbers have been removed.

The concepts of controller and processor were further explained. A controller is an entity that decides why and how data is processed, while a processor processes data on behalf of the controller. It was highlighted that controllers and processors are obligated to protect data under the GDPR regardless of whether they are public or private entities.

The chapter also presented the territorial scope of the GDPR, which as a general rule, is determined by the location at which the controller or processor is established and not by the place at which the data in question is processed. Therefore, a controller or processor established in the EU cannot avoid responsibility under the GDPR simply by locating its processing activities outside the EU. The principles in Article 5 GDPR, which are applicable in themselves and serve as guidelines for the interpretation of other provisions, were also briefly explained.

The definition of "data concerning health" in Article 4(15) GDPR was clarified. Health data within the meaning of the GDPR includes all personal data that relates to a person's physical or mental health, including data that relates to the provision of health care services, which reveal information about a person's health status. The WP29 has provided guidance on to what extent data generated by apps and devices can fall under the definition of health data.

Two conclusions that can be drawn from the WP29's guidance are that data generated within a medical context can be presumed to fall under the GDPR's definition of health data, while data generated by lifestyle apps and devices to a greater extent need to be determined on a case-by-case basis. This leaves a lot of room for the controller to assess whether the data collected constitutes health data, or if it might in the future, when combined with other data or if measured over time. While non-commercial research entities such as university hospitals may to a greater extent conduct research involving data that can be presumed to fall under the definition of health data in the GDPR, for example, information collected through the provision of health services, companies are likely to face more complex situations when assessing whether data generated by, for example, app usage and social media constitute health data

within the meaning of the GDPR. It has also been highlighted by scholars that big data analytics could further complicate this assessment, as it can derive sensitive data from "regular" data.

While it is not always clear whether or not data falls under the GDPR's definition of health data, this uncertainty may not cause significant problems in practice. Companies that offer devices and apps that, for example, aim to work as a tool for people to monitor their own health, as well as stakeholders looking to apply AI within the area of healthcare, would be wise to assume that at one point or another they will process health data and take this into account from the beginning. In the following chapters, when the thesis refers to "health data" it should be understood as "data concerning health" within the meaning of Article 4(15) GDPR.

Lastly, it was mentioned that the processing of special categories of data, such as health data, is generally prohibited, but allowed when processed for scientific research purposes. In the next chapter, the scientific research exemption in Article 9(2)(j) GDPR and the debate regarding its breadth will be explained further.

3 A Guide to the Debate on the Extent of the Scientific Research Exemption

The aim of this chapter is to outline the applicability of the scientific research exemption in Article 9(2)(j) GDPR. The chapter will also describe the debate on the extent of the scientific research exemption in relation to data processing by private companies. It should again be mentioned that although there is a discussion on the extent of it, it has mainly been centered around whether private companies should benefit from it or not.

3.1 Scientific Research Purposes

According to Article 9(2)(j) GDPR, the processing of special categories of personal data is allowed if:

> the processing is necessary for archiving purposes in the public interest, scientific or historical research purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and interest of the data subject.

The GDPR does not contain a binding definition of "scientific research purposes." Instead, guidance can be found in Recital 159 of the GDPR, according to which "scientific research purposes should be interpreted in a broad manner including for example technological development and demonstration, fundamental research, applied research and privately funded research." Furthermore, the term "should also include studies conducted in the public interest in the area of public health." Recital 159 further states that the application of the provision on scientific research purposes should take into account the Union's objective under Article 179(1) TFEU of achieving a European research area.

In relation to the call for a broad interpretation of the scientific research purposes that the recital calls for, the WP29 has taken the position that "the notion may not be stretched beyond its common meaning" and understands it as "a research project set up in accordance with relevant sector-related methodological and ethical standards, in conformity with good practice."[82] This view was later endorsed by the EDPB.[83]

[82] WP29, "Guidelines on Consent under Regulation 2016/679" (2018), p. 27-28.
[83] EDPB (2018), p. 1.

The European Data Protection Supervisor (EDPS) has also, in a preliminary opinion, issued a statement on what activities can be considered as "scientific research."[84] The EDPS notes that scientific research can be conducted not only by academic researchers but also by commercial companies, governmental institutions and non-profit organizations.[85] However, for a controller to simply claim that it is carrying out scientific research is not sufficient to invoke the exemption in the GDPR. For the scientific research exemption to apply, the EDPS states that not only must "relevant sectoral standards of methodology and ethics apply" – which is also the view of the EDPB – but it must also be conducted "with the aim of growing society's collective knowledge and wellbeing, as opposed to serving primarily one or several private interests."[86]

The GDPR as a regulation is binding in its entirety and directly applicable in all Member States.[87] However, the reference to Member State law in Article 9(2)(j) GDPR is not completely clear. According to Recital 8, where the GDPR provides for specifications or restrictions of its rules by Member State law, Member States may incorporate elements of the regulation into their national law. Recital 10 states that the GDPR provides a margin of maneuver for Member States to specify its rules, including the processing of special categories of data. Furthermore, Recital 52 states that derogating from the prohibition on processing special categories of personal data should be allowed when provided for in Union or Member State law.

3.2 The Extent of the Scientific Research Exemption

The breadth of the scientific research exemption has been widely discussed by scholars. Pormeister states that the research exemption will in practice cover all types of research, including commercial research. Furthermore, Pormeister argues that the definition of research in the GDPR could expand the scope of Member State law if it was to interpret research more narrowly than the GDPR.[88] Ducato states that the notion of scientific research in Recital 159 appears to include activities conducted for profit, for example, "experimental development carried out by a company to improve or offer new services."[89] Wiese Svanberg writes that the GDPR makes no distinction between

[84] The European Data Protection Supervisor (EDPS) is an independent EU authority which, under article 58(3)(c) of Regulation 2018/1725, is empowered "to issue on his or her own initiative or on request, opinions to Union institutions and bodies and to the public on any issue related to the protection of personal data."

[85] EDPS (2020), p. 11.

[86] Ibid., p. 11-12.

[87] Article 288 TFEU.

[88] Pormeister (2017), p. 138, 140. As stated in section 1.4, Pormeister argues that the broad exemption shifts the balance of interests rather heavily in favor of the data processor/controller, which leaves the data subject with limited control over the processing of their genetic data for scientific research (see Pormeister [2017], p. 146).

[89] Ducato (2020), p. 3.

scientific research conducted in the public interest and scientific research conducted for private or purely commercial interests. Provided that the provisions of the GDPR and relevant Member State law are adhered to, purely private or commercial interests may be pursued through processing personal data for scientific research purposes.[90]

Slokenberga has criticized the abovementioned view of the EDPS, that research must be conducted "with the aim of growing society's collective knowledge and wellbeing, as opposed to serving primarily one or several private interests." She states that there are several problems with the distinction between "collective knowledge" and "primarily one or several private interests," as it fails to take into account the reality in which scientific research is carried out and commercialization as a driver of scientific advances. However, she notes that defining scientific research should not result in depriving data subjects of their right to protection and give researchers carte blanche for their activities, if there is no benefit to society that comes with it. In that case, one may agree with the EDPS that research should bring some kind of value to the public, she states. Furthermore, Slokenberga concludes that although the view of the EDPS reflects the approach of the CJEU of interpreting exceptions narrowly, it does not align with the legislator's intention to interpret it broadly.[91]

Mészáros and Ho argue that commercial AI research should not benefit from the research exemption without public interest and similar safeguards as academic research.[92] They emphasize that it would be essential to differentiate between academic and commercial research, as is done in the EU Copyright Directive, which defines "research organization" as an entity that conducts scientific research "on a not-for-profit basis or by reinvesting all the profits in its scientific research" or "pursuant to a public interest mission."[93] Bentzen states that failing to define "scientific research" in the GDPR may extend the privilege it confers on such research to an unintentionally wide range of actors and activities, which may pose risks to the fundamental rights of research participants.[94]

It should be noted that public interest is, just as scientific research purposes, a concept of EU law and is not defined in the GDPR. However, Recitals 45 and 46 provide some examples of what can be considered public interest, which are "public health and social protection and the management of health care services" and "monitoring epidemics and their spread or in situations of humanitarian emergencies, in particular in situations of natural and man-made

[90] Wiese Svanberg (2020), p. 1249.
[91] Slokenberga (2021), p. 24-25.
[92] Mészáros and Ho (2021), p. 2-3.
[93] Ibid., p. 8; Article 2(1) Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC.
[94] Bentzen (2020), p. 344.

disasters." Linguistically, "public interest" can be assumed to refer to something that is of interest or concerns people on a broader scale, as opposed to a specific or individual interest.[95]

The question of whether invoking the scientific research exemption should require public interest has also been discussed by others. Verhenneman states that scientific research must be assumed to bring some value to society and further, "while that value to society does undoubtedly not require the activity to be of public interest, it does require that society enjoys at least some of the benefits."[96] Wouters et al. state that while the legislator encourages the development of a European research area, Recital 53 also requires that scientific research for health-related purposes is pursued with a public interest objective, which they conclude limits the types of institutions that conduct scientific research.[97] Kruus has pointed out the fact that the legislator has explicitly set out a requirement for archiving purposes to be in the public interest, but has not stated the same for scientific or historical research purposes or statistical purposes.[98]

To conclude, the discussion on the scientific research exemption can perhaps best be summarized the following way, until the EU legislator or the CJEU provides clarification, the term scientific research in the GDPR remains a grey area.[99]

3.3 Summary and Concluding Remarks

This chapter explained the applicability of the scientific research exemption in Article 9(2)(j) GDPR and outlined the ongoing discussion regarding the extent of it.

First of all, the reference to Member State law in Article 9(2)(j) GDPR is not entirely clear and the recitals appear to contain conflicting messages. Recital 52 could be interpreted as requiring that an exemption from the general prohibition on the processing of special categories of personal data must be laid down in national law, while other recitals seem to suggest that Member States may further specify the conditions for the processing of sensitive data. However, as the term "scientific research purposes" is a concept of EU law, the

[95] Öman (2023), legal commentary on Article 6(1)(e) GDPR.

[96] Verhenneman (2021), p. 297.

[97] Wouters et al. (2021), p. 207. Recital 53 GDPR: "Special categories of personal data which merit higher protection should be processed for health-related purposes only where necessary to achieve those purposes for the benefit of natural persons and society as a whole […] or for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, based on Union or Member State law which has to meet an objective of public interest, as well as for studies conducted in the public interest in the area of public health."

[98] Kruus (2023), p. 66.

[99] Ibid., p. 67; see also Bentzen (2020), p. 349.

reference to Member State law should not be interpreted as requiring Member States to define "scientific research purposes," but rather to provide for the conditions under which it is allowed. The interpretation of the reference to Member State law in Article 9(2)(j) GDPR in relation to Swedish law will be further discussed in section 4.2.

It should be mentioned again that the GDPR does not contain a binding definition of what is covered by "scientific research purposes." However, Recital 159 GDPR calls for a broad interpretation of the term. The WP29 and the EDPS have put forward some limitations to this definition, which is that research must be conducted within a sectoral framework, consisting of methodological and ethical standards. This is a fair delimitation given that, although the GDPR calls for a broad interpretation of scientific research purposes, it should not mean that controllers are free to define it as they see fit. As expressed by Slokenberga, processing under the scientific research exemption does not give controllers carte blanche for their processing activities.

In addition to the fact that research should be conducted in accordance with methodological and ethical standards, the EDPS notes that research must be conducted with the aim of increasing society's collective knowledge, as opposed to serving primarily private interests. Given that it is a preliminary opinion, it should be highlighted that the statement cannot be assumed to carry too much bearing for the interpretation of the term. Furthermore, besides the fact that it, as Slokenberga argues, fails to consider the reality in which research is conducted, it lacks support in the GDPR, which acknowledges that the term should include privately funded research. While privately funded research does not automatically entail that research is also conducted with a private interest, the inclusion of it can be interpreted as an acknowledgment by the EU legislator that research can be conducted with a private interest, and that private and public interests can be intertwined in research. Furthermore, it would be difficult to achieve the objective under Article 179(1) TFEU, referred to in Recital 159, of encouraging the European Research Area to become more competitive, if the scientific research exemption excludes research serving private interests. It also risks excluding research projects which, although they may be conducted on a for-profit basis, could have a positive impact on people's health.

As mentioned above, there is a discussion about whether processing under the scientific research exemption requires public interest. It was highlighted by Kruus, that the EU legislator has explicitly stated that archiving purposes, to be covered by the exemption in Article 9(2)(j), must be in the public interest, while the same requirement has not been set forward for scientific research. Recital 159 states that scientific research purposes should include studies conducted in the public interest in the area of public health but does not indicate that the requirement of public interest applies to all research. However, the meaning of Recital 53 GDPR is not entirely clear. According to Recital 53,

scientific research conducted for "health-related purposes" appears to require Union or Member State law to meet an objective of public interest.

The term "health-related purposes" is vague, and it is unclear whether research related to, for example, the integration of AI in mobiles apps for diet monitoring could be covered by the term or if it is more focused on research related to the provision of health care services, such as cancer treatment research. It probably covers more than the term public health, as Recital 53 also specifically mentions "studies conducted in the public interest in the area of public health." In addition, it is also unclear whether Member States would then be required to regulate research for health-related purposes separately. The requirement that Member State law has to meet an objective of public interest when scientific research is conducted for "health-related purposes," is, however, set out in the non-binding part of the GDPR. Therefore, in the absence of more clear guidance on what is covered by the term "health-related purposes," processing under the scientific research exemption cannot be considered to require public interest, as it is not stipulated in Article 9(2)(j).

The general view among scholars also appears to be that public interest is not necessarily a requirement set out by the GDPR for scientific research, but rather that processing under the scientific research exemption should bring some form of societal value, in the sense that the research benefits people rather than just generating profit. Whether companies engaging in data-driven research can argue that the research has a societal value may vary. For companies engaging in research where the aim is to develop devices that can be used by physicians in the provision of healthcare, for example, an AI tool that can detect abnormalities in X-rays, it is probably fairly easy to argue that there is a societal value. However, for companies engaging in research that results in, for example, a commercialized mobile app containing an AI solution to improve the self-management of people's health, it may depend on factors such as how many similar apps exist, whether it can have a positive impact on people's health status, or whether it is rather a way for the company to generate profit.

In order to answer the first research question concerning the extent to which Swedish companies can claim the scientific research exemption in Article 9(2)(j) GDPR when processing health data, Chapter 4 will examine Swedish law. However, a couple of conclusions can be drawn which will form the basis for the analysis of the scientific research exemption in relation to Swedish law. Firstly, the GDPR does not differentiate between research conducted by public or private entities or whether research is conducted with a commercial interest or not. Furthermore, the inclusion of "privately funded research" can be interpreted as an acknowledgment by the EU legislator that several interests may interplay in research. Secondly, as stated by the WP29 and later endorsed by the EDPB, the term "scientific research" should be understood as a project that adheres to methodological and ethical standards.

Although the term should be interpreted broadly, controllers cannot define it according to their discretion. Lastly, processing under the scientific research exemption does not require public interest, understood as a concept of EU law. However, the general understanding in literature appears to be that scientific research should benefit people or society, rather than solely focusing on profit. Whether companies engaging in data-driven research can argue that there is a societal value may vary. The next chapter will examine the "scientific research exemption" in relation to Swedish law.

4 The Scientific Research Exemption and Swedish Law

In Swedish law, until the entry into force of the GDPR, public and private research entities processed personal data primarily under the Personal Data Act, which was the Swedish implementation of Directive 95/46/EC.[100] With the entry into force of the GDPR and the Swedish Act Containing Supplementary Provisions to the EU General Data Protection Regulation, the Directive and the Personal Data Act were repealed.[101] The changes that were made to various laws as a result of the GDPR intended to ensure the continued processing of data for scientific research purposes, while still protecting the privacy of individuals, regardless of whether the processing is carried out by public or private researchers.[102]

When research involving the processing of special categories of personal data is carried out in Sweden, the Swedish Ethical Review Act applies alongside the GDPR.[103] The Ethical Review Act contains its own definition of research and the work on defining research in Swedish law has mainly taken place within the framework of the ethical review system.[104] The aim of this chapter is to examine the relationship between "scientific research purposes" in the GDPR and "research" as defined in the Ethical Review Act, in order to discuss the effect it may have on the ability of Swedish companies to claim the scientific research exemption in Article 9(2)(j) GDPR when processing health data.

4.1 The Ethical Review Act

The purpose of the Ethical Review Act is to protect the individual and the respect for human dignity in research.[105] Research falling within the scope of the law must therefore undergo an ethical review and can only proceed after approval, which pertains to a specific project or part of a project.[106] In order to reuse data from a prior study in a different research project, approval under the Ethical Review Act has to be obtained again.[107] Given that an approval only relates to a specific project or part of a project, it is not possible to obtain

[100] Prop. 2017/18:298, p. 18.
[101] Ibid., p. 19.
[102] Ibid., p. 1.
[103] Section 3 of the Ethical Review Act.
[104] SOU 2017:50, p. 90.
[105] Section 1 of the Ethical Review Act
[106] Section 6 of the Ethical Review Act.
[107] Swedish Ethical Review Authority (2023), p. 42.

for example approval to conduct research within a certain field "in the foreseeable future."[108]

According to Section 9 of the Ethical Review Act, research can only be approved if the risks it may entail for the health, safety, and personal integrity of the research subjects are outweighed by its scientific value. Furthermore, when research involves the processing of sensitive data, the processing itself needs to be approved through an ethical review.[109] Such processing should only be approved if it is necessary for the performance of the research.[110] A general requirement for research involving humans is that it carries a theoretical and/or practical value to society at large and is expected to generate important knowledge. This may involve the collection of personal data to identify correlations, or medical research that can contribute to improved diagnosis, treatment, or preventive measures in health care. Evaluating the scientific value is essential for the subsequent risk assessment.[111] Furthermore, research can only be approved if it is carried out by or under the supervision of a researcher with the necessary scientific competence, for example, a researcher with a doctorate.[112]

The Ethical Review Act requires in certain cases that research participants have consented to and are provided with information about the research project, such as what research methods will be used. This does not apply to research involving the processing of sensitive data.[113] However, the Swedish Ethical Review Authority has pointed out in its guidelines that consent to participate in a scientific study is a principle of research ethics and that people recruited to participate in a research project should always be informed of what their participation entails. However, the Authority notes that some flexibility applies to the requirement of consent when the research involves the processing of sensitive data. For example, if data is collected through a questionnaire, the person filling out the form can be assumed to have provided their consent, if the participants can be considered to have received adequate information about the research project.[114]

Applications for ethical review of research are examined by the Swedish Ethical Review Authority. In its guidelines, a number of documents that are required to be enclosed with the application are listed. These include, for example, a research plan, information to be provided to research subjects, a list of variables when requesting data from existing registers, and the CV of the researcher in charge.[115] It is possible to obtain approval for several research

[108] Prop. 2002/03:50, p. 195.
[109] Section 6 of the Ethical Review Act.
[110] Section 10 of the Ethical Review Act-
[111] Prop. 2002/03:50, p. 98-99.
[112] Section 11 of the Ethical Review Act; prop. 2002/03:50, p. 100.
[113] Section 13 of the Ethical Review Act interpreted a contrario.
[114] Swedish Ethical Review Authority (2023), p. 41.
[115] Swedish Ethical Review Authority (2023), p. 107.

projects described in one and the same application if they have a clear connection. However, it is insufficient for a researcher to simply wish to study certain material from many different perspectives.[116]

If an application is denied, the activities described in the application are not allowed to be carried out. The Ethical Review Authority states that the most common reasons for denial are that the risk or burden that the research entails for the research subjects is not outweighed by its scientific value, that the research questions are not clear enough, or that it is unclear how the questions can be answered with the method set forward in the application. An application can also be rejected. One reason for this is that the project does not meet the definition of research in the Ethical Review Act. If an application is rejected because of this, the activities are not allowed to be carried out on the legal basis that applies to scientific research, i.e. the scientific research exemption.[117]

4.2 Ethical Review as a Safeguard under the GDPR

As explained in section 3.1, the scientific research exemption under Article 9(2)(j) GDPR allows for the processing of special categories of personal data for scientific research purposes "in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject." The reference in Article 9(2)(j) GDPR to Member State law, which is not completely clear, has been discussed in the preparatory works to the Swedish Act Containing Supplementary Provisions to the EU General Data Protection Regulation.[118] The Swedish government has interpreted the reference to Member State law as not requiring the implementation of the scientific research exemption in itself in national law, but that the GDPR requires some additional form of support in Swedish law.[119]

Research involving the processing of special categories of personal data is therefore subject to ethical review pursuant to Section 3 of the Ethical Review Act. Ethical review has been deemed to constitute a suitable and specific measure under Swedish law that is required in order to process sensitive data for scientific research purposes under Article 9(2)(j) GDPR.[120] In the preparatory work relating to the entry into force of the GDPR, it was discussed whether a "research data law" should be introduced, and if so, whether it should include a provision that specified that sensitive personal data could be processed under Article 9(2)(j) GDPR if the processing had been approved in

[116] Swedish Ethical Review Authority (2023), p. 109.
[117] Ibid., p. 117-118.
[118] See prop. 2017/18:105, p. 75-75 and SOU 2017:39, p. 162 ff.
[119] Prop. 2017/18:105, p. 75.
[120] Prop. 2017/18:298, p. 84, 88.

accordance with the Ethical Review Act. Such a provision was however considered to be redundant, since the scientific research example in itself did not have to be implemented in Swedish law and since the regulation in the Ethical Review Act was considered sufficient to meet the requirements in Article 9(2)(j).[121]

Another question that was discussed in the preparatory works was whether all research involving the processing of personal data, and not only the processing of sensitive data, should be subject to ethical review.[122] Ethical review would then constitute a safeguard under Article 89(1) GDPR, which requires that processing for scientific research purposes is subject to "appropriate safeguards." However, expanding the scope to include all personal data, rather than just sensitive data, was considered too excessive and would risk undermining the ethical review system. If there is no risk of harm, an ethical review should not be carried out.[123] Processing of personal data for scientific research purposes, where no sensitive data is involved, is therefore not subject to ethical review.

4.3 Territorial Scope of the Ethical Review Act

The territorial scope of the Ethical Review Act is specified in Section 5 of the Ethical Review Act: It applies to research conducted in Sweden. If a Swedish research principal (i.e. a public authority or a natural or legal person in whose establishment the research is conducted) participates in an international research project and part of the research is carried out in Sweden, said part is subject to ethical review.[124]

The territorial scope of the Ethical Review Act does not align with the territorial scope of the GDPR, which as a general rule is decided by where the controller or processor is established, regardless of where the processing takes place. In a situation where research involving the processing of, for example, health data is conducted outside Sweden, but by a controller established in Sweden, the Ethical Review Act will not apply. However, the controller is still bound by the GDPR's prohibition of processing special categories of personal data. In that case, the preparatory work refers to the option of obtaining consent for the processing.[125]

The Swedish Ethical Review Authority clarifies in its guidelines that when a research principal in Sweden interacts with research participants via digital

[121] Prop. 2017/18:298, p. 90; Moreover, a "research data law" was never introduced, in part because it would contain only a small number of provisions, see prop. 2017/18:298, p. 139.
[122] Ibid., p. 81.
[123] Ibid., p. 82.
[124] Prop. 2002/03:50, p. 109, 194.
[125] Prop. 2017/18:298, p. 90, 139.

media, the research is deemed to be conducted in Sweden, even if the participants are located elsewhere. The fact that the research can be considered to be conducted in another country at the same time does not affect the requirement of ethical review under the Ethical Review Act. Furthermore, research is considered to be carried out in Sweden when the research principal is located outside of Sweden, but interacts with research participants located in Sweden via digital channels.[126]

4.4 Research as Defined in the Ethical Review Act

The Ethical Review Act defines research as:

> Scientific experimental or theoretical work or observational research studies, if the work or studies are carried out to acquire new knowledge, and/or development work on a scientific basis, but not such work or studies that are performed solely within the framework of higher education at basic or advanced level.[127]

Scientific experimental or theoretical work and development work on a scientific basis refers to both basic and applied research.[128] Basic research is work undertaken without any particular application or use in mind, while applied research is mainly focused towards a specific, practical purpose or objective.[129] "Development work on a scientific basis" refers to the "imaginative" and "systematic" use of scientific knowledge and other types of information to achieve new products, new processes, new systems, or significant improvements to existing systems.[130] The research should aim at acquiring new knowledge, which includes research that repeats previously conducted research in order to reinforce or challenge previous findings and conclusions. The criterion of "scientific" entails that the work should be part of a process "where knowledge is systematized and structured through theoretical developments and the application of methodological tools."[131] By emphasizing the scientific approach both in the acquisition of new knowledge and in development work, research is distinguished from other activities that may be of a similar nature, such as quality assurance, performance monitoring, or journalism.[132]

The Swedish government has made the assessment that the definition of research in the Ethical Review Act is covered by the term scientific research

[126] Swedish Ethical Review Authority (2023), p. 62
[127] Section 2 of the Ethical Review Act.
[128] Prop. 2007/08:44, p. 50.
[129] OECD (2015), p. 45.
[130] Prop. 2002/03:50, p. 192.
[131] Prop. 2007/08:44, p. 50.
[132] Ibid., p. 19.

purposes in the GDPR.[133] However, there might be situations where the definition of research in the Ethical Review Act does not fully cover what is considered scientific research purposes under the GDPR. The government has not been able to ascertain what situations that may be and has instead left it to be determined through the application of the law. As regards the examples in Recital 159 of what activities should be included by scientific research purposes, the preparatory work concludes that for example fundamental research, applied research, and privately funded research are clearly covered by the definition of research in the Ethical Review Act. The same applies for studies conducted in the public interest in the area of public health. As for technological development and demonstration, such activities may be classified as development work on a scientific basis.[134]

In a government inquiry, a committee emphasized that definitions of research in national law do not affect the assessment of whether or not certain processing is done for scientific research purposes under the GDPR.[135] Furthermore, in another government inquiry, the committee observed that Recital 159 does not include any requirements regarding the qualifications of the researchers or the organizational framework in which the research should take place, apart from the mention of privately funded research. The committee concluded that it is therefore the content of the activity that is decisive for the assessment of whether or not the activity falls under scientific research purposes.[136]

However, external factors may affect the assessment of what qualifies research under the Ethical Review Act. The Swedish Ethical Review Authority states in its guidance on ethical review that if there is an intention to publish the research results, it demonstrates an expectation to acquire new knowledge, which it concludes is also an integral part of the definition of research in the Ethical Review Act. While there may be valid reasons for not publishing the results immediately, a lack of intention within a project to publish the results may indicate that the work does not constitute scientific research, even if it is carried out by researchers through scientific questions and methods. Therefore, as a general rule, if the purpose of the work is to only make the results available within, for example, a company, it typically does not constitute research within the meaning of the law.[137]

The Swedish Research Council has also stated in their guidance on what constitutes good research practice, that researchers generally have an obligation to publish their results.[138] Furthermore, practice developed by the Ethics Review Appeals Board, which reviews appealed decisions by the Swedish

[133] Prop. 2018/19:165, p. 19.
[134] Prop. 2017/18:298, p. 134.
[135] SOU 2018:36, p. 74.
[136] SOU 2017:50, p. 93-94.
[137] Swedish Ethical Review Authority (2023), p. 74.
[138] Swedish Research Council (2017), p. 52

Ethical Review Authority, demonstrates that factors such as if the research is conducted by a person with scientific expertise or if there is an intent to publish the research in, for example, a peer-reviewed journal, indicate that the activity constitutes research.[139]

4.5 Summary and Concluding Remarks

This chapter examined the processing of special categories of personal data for scientific research purposes in Swedish law. As Article 9(2)(j) GDPR refers to Member State law, different rules for the processing of data for scientific research may apply depending on the Member State in question. In order to answer the first research question of to what extent Swedish companies can claim the "scientific research exemption" in Article 9(2)(j) GDPR when processing health data, Swedish law is here used as a practical example.

First of all, it should be emphasized that "scientific research purposes" is a concept of EU law, which means that Member States cannot provide their own definitions of it. However, because of the reference to Member State law, national law may provide additional conditions under which processing for scientific research purposes is allowed. In Swedish law, ethical review under the Swedish Ethical Review Act has been deemed to constitute a suitable and specific measure that is required in order to process sensitive data for scientific research purposes under Article 9(2)(j) GDPR. As mentioned in Section 3.2, it has been argued that the definition of scientific research in the GDPR could expand the scope of Member State law if it was to interpret research more narrowly than the GDPR. While that may be the case in theory, Swedish law could in practice limit the scope of the scientific research exemption, if the processing activities of a company fall outside the definition of research in the Ethical Review Act, since ethical review is a condition for processing data under Article 9(2)(j) GDPR. Therefore, the definition of research in the Ethical Review may affect the ability of companies to claim the scientific research exemption in the GDPR.

Recalling the first conclusion from the previous chapter, the GDPR does not differentiate between research conducted by public or private entities or whether research is conducted with a commercial interest or not. The definition of research in the Ethical Review Act also does not distinguish between public and private researchers. The provision, however, emphasizes that the work should aim at acquiring new knowledge, which the preparatory work highlights also applies to development work on a scientific basis. Furthermore, the preparatory work states that a general requirement for research involving humans is that it carries a theoretical and/or practical value to society at large and is expected to generate important knowledge. While the focus on acquiring new knowledge and the theoretical and/or practical value that the

[139] Swedish Ethical Review Authority (2023), p. 73-74.

research must carry does not rule out activities that are conducted with commercial interests, it may be difficult for companies to argue that activities that are undertaken with solely commercial interests aim at acquiring new knowledge.

Recalling the second conclusion, research under the scientific research exemption should adhere to sector-related methodological and ethical standards. While this is not stated in the provision or the recital concerning scientific research, it has been argued by the WP29 and later endorsed by the EDPB. The thesis argues that this is a fair delimitation of the otherwise broad interpretation of scientific research, as the exemption should not be intended to ease the requirements of the GDPR for companies in their profit-making activities, but rather offer a favorable regime in order to allow for scientific advances. Ethical review as a suitable and specific safeguard therefore ensures that companies adhere to methodological and ethical standards, which can be especially important in relation to data-driven research focused on the healthcare sector, for example, to ensure safety and avoid algorithm bias.

The third conclusion that was drawn in the previous chapter was that the GDPR does not require public interest, but that the general opinion among scholars is that research under the scientific research exemption should benefit people or society, rather than solely focusing on profit. The emphasis in the Ethical Review Act that the work should aim at acquiring new knowledge and the general requirement for research involving humans of carrying a theoretical and/or practical value to society at large, resembles a requirement of public interest – not necessarily in the strict sense that Recitals 45 and 46 provide examples of, but rather that research is of interest to or concerns people on a larger scale.

Part of the criticism that has been directed at commercial research, for example by Mészáros and Ho, who have argued that commercial research should not benefit from the scientific research exemption without public interest and similar safeguards as academic research, is therefore addressed in Swedish law. However, while research as defined in the Ethical Review Act does not rule out commercial interests, the threshold for what constitutes research is higher than what is set forward by Recital 159, which does not require public interest.

The stricter requirement of what constitutes research set forth by the Ethical Review Act can be illustrated with a comparison between technological development in Recital 159 and development work on a scientific basis in the definition of research in the Ethical Review Act. For example, the processing activities by a company to improve or offer new services may count as technological development in the GDPR. However, the activities may not classify as development work on a scientific basis in the Ethical Review Act, as they may not reach the threshold of “significant improvement to an existing

system." This is a clear disadvantage for companies engaging in data-driven research as the ability to improve systems could be an integral part of their activities. As stated above, the purpose of the scientific research exemption should not be to ease the requirements of the GDPR for companies in their profit-making activities, but rather offer a favorable regime in order to allow for scientific advances. Therefore, the companies that benefit from the scientific research exemption when carrying out research in Sweden are the companies whose processing activities also carry a scientific value and therefore meet the definition of research in the Ethical Review Act, which is where the major benefits for the healthcare sector are found.

However, two barriers in regard to the requirement of ethical review pursuant to the Ethical Review Act are identified below, which may have an inhibiting effect on the ability of companies to invoke the scientific research exemption. First of all, one barrier to the ability of companies to claim the scientific research exemption in the GDPR can be attributed to the territorial scope of the Ethical Review Act. As outlined above, the Ethical Review Act only applies to research conducted in Sweden. Therefore, if research involving the processing of health data is conducted outside of Sweden, but by a controller established in Sweden, the Ethical Review Act will not apply.

For example, consider a situation where a tech company established in Sweden has located its processing activities outside the EU. Since the processing is carried out in the context of the activities of the tech company in Sweden, it is bound by the GDPR according to Article 3(1) and therefore also by the general prohibition on the processing of sensitive data in Article 9(1). The company wants to develop a health application through the processing of health data, which falls under technological development in Recital 159. The activities also fall under the definition of research in the Ethical Review Act, namely development work on a scientific basis, as the company uses scientific knowledge to achieve a new product. However, the company is not able to claim the scientific research exemption in the GDPR, as the research is not conducted in Sweden and falls outside the territorial scope of the Ethical Review Act. Since the processing involves health data, the company must find another exemption, such as consent, which may prove difficult if the activities involve data from a large number of people. While the barrier in terms of the territorial scope applies to both public and private researchers, it could affect companies to a greater extent, as researchers such as public universities are probably more likely to undertake research activities in the country where they are established.

Another barrier is that approval under the Ethical Review Act only pertains to a specific project or part of a project, which removes some of the flexibility offered by the scientific research exemption in the GDPR. The requirement to obtain an ethical review for every research project constitutes a barrier for companies whose processing activities routinely fall under scientific research

purposes in the GDPR. This may prove to be a burdensome requirement for startups, for example, that are looking to engage in data-driven research, but may not have the resources to allocate time and gather material to go forward with an ethical review application.

The next chapter will outline the legal grounds on which the processing of health data must be based, apart from being covered by an exception in Article 9(2) GDPR.

5 Legal Bases for the Processing of Health Data for Scientific Research Purposes

As explained in section 2.3, personal data must be processed lawfully, fairly, and in a transparent manner in relation to the data subject, according to the principle in Article 5(1)(a).[140] The requirement of lawfulness of processing is further specified in Article 6 GDPR which states that processing is lawful only if and to the extent that it is based on at least one of the six bases stipulated in Article 6. The legal grounds are:

- Consent of the data subject;
- Contract and pre-contractual relationship;
- Processing for compliance with a legal obligation to which the controller is subject;
- Processing is necessary in order to protect the vital interests of the data subject or of another natural person;
- Processing for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;
- Processing on grounds of legitimate interests pursued by the controller or by a third party.[141]

In the private sector, consent may be of utmost importance for allowing data processing in cases where there is no contractual obligation, no detailed rules on the appropriate legal basis or where it is particularly difficult to assess the scope of the "legitimate interests of the controller or of a third party."[142] It is important for researchers to carefully consider the appropriate legal base depending on the specific context, because of the conditions and opportunities that come with it.[143] Below, four of the legal bases will be explained further, as they may be relevant in relation to companies engaging in data-driven research. These are: consent, contract and pre-contractual relationship, task carried out in the public interest, and legitimate interests pursued by the controller or by a third party

[140] Article 5(1)(a) GDPR.
[141] Article 6(1)(a)-(f) GDPR.
[142] Kotschy (2020), p. 329.
[143] Quinn and Quinn (2018), p. 1011.

5.1 Consent of the Data Subject

One of the grounds for lawful processing of personal data is whether the data subject has consented to it.[144] Consent is defined as any freely given, specific, informed, and unambiguous indication that the data subject agrees to the processing of his or her personal data and can be given by, for example, a written statement, including electronically (e.g. by ticking a box on a website), or an oral statement.[145] However, silence, pre-ticked boxes, or inactivity does not constitute consent.[146] In the case that a declaration of consent is already formulated by the controller, it should be provided in an easily understandable and accessible form, using clear and plain language, and should not include any unfair terms.[147] The data subject has the right to withdraw his or her consent at any time.[148] If the controller has another legal basis for the processing, besides consent, the controller can continue the processing.[149] However, as has been stated by the WP29, is it not possible for a controller to swap from consent to another legal basis if the controller encounters problems in relation to the consent, unless that other legal basis has been specified before the collection of data.[150]

For consent to be freely given, there must not be a clear imbalance between the data subject and controller, in particular where the controller is a public authority. In such circumstances, consent is unlikely to be freely given.[151] Another example is the power balance between an employer and an employee, as the dependency of this relationship is likely to result in consent not being freely given.[152] If a data subject does not have a genuine choice, feels compelled, or faces repercussions if they do not consent, it cannot be considered freely given.[153] The Swedish government has stated that while private researchers must always determine whether or not there may be an imbalance between them and the data subject, as a general rule, private researchers should not be hindered from using consent for scientific research purposes.[154]

As regards the processing for scientific research purposes, Recital 33 is of particular interest, stating that it is often not possible to fully identify the purpose of the processing for scientific research at the time of data collection. Data subjects should therefore be allowed to give their consent to "certain areas of scientific research," as long as it is in accordance with recognized

[144] Article 6(1)(a) GDPR.
[145] Article 4(11) and Recital 32 GDPR.
[146] Recital 32 GDPR.
[147] Recital 42 GDPR.
[148] Article 7(3) GDPR.
[149] Kosta (2020), p. 351.
[150] WP29 (2018), p. 23
[151] Recital 43 GDPR.
[152] EDPB (2020), p. 9
[153] Ibid., p. 7
[154] Prop. 2017/18:298, p. 42.

ethical standards for scientific research. The WP29 has noted that Recital 33 allows for the purpose to be described at a more general level if it is not possible to specify the purpose within a scientific research project at the outset. However, it further states that in view of the special regulation of processing of special categories of personal data, the "flexible approach of Recital 33 will be subject to a stricter interpretation and requires a high degree of scrutiny."[155]

The use of consent as a legal basis may not always be the most reliable option, which can be particularly evident in research involving big data.[156] One challenge in particular concerning big data analytics is that the value that the personal data holds is not always apparent at the time of data collection, when consent is usually given. If future uses have not been articulated at this time, it may require controllers to go back to the data subjects for their amended consent, which might prove too costly to undertake, even if these future uses may hold significant value to both individuals and society at large. Cate and Mayer-Schönberger argue that what used to be a straightforward relationship between controllers and data subjects has become complicated because of the combination of data sets in big data analytics and as the processor can quickly change, which could make it hard for individuals to fully understand the complexity of the situation they are asked to consent to.[157]

Besides using consent as a ground for lawful processing on a general basis in accordance with Article 6(1)(a), consent can also be used to allow the processing of health data under Article 9(2)(a), however, the consent has to be "explicit," which is a higher threshold than what is required by Article 6(1)(a).[158]

5.2 Necessary for the Performance of a Contract

When processing is necessary for the performance of a contract to which the data subject is a party or in order to take steps at the request of the data subject prior to entering into a contract, it can be based on Article 6(1)(b) GDPR. An example of processing that is necessary for the performance of a contract is the processing of a data subject's address in order to deliver goods purchased online.[159]

[155] WP29, "Guidelines on Consent under Regulation 2016/679" (2018), p. 28.
[156] Ducato (2020), p. 7; Quinn and Quinn (2018), p. 1013.
[157] Cate and Mayer-Schönberger (2013), p. 67-68.
[158] Georgieva and Kuner (2020), p. 377.
[159] EDPB, "Guidelines 2/2019 on the Processing of Personal Data under Article 6(1)(b) GDPR in the Context of the Provision of Online Services to Data Subject" (2019) p. 10.

A controller that seeks to base its processing on this ground must ascertain that the processing is "objectively necessary."[160] In this regard, the EDPB has endorsed the view of the WP29, which has stated that the criteria of necessity "must be interpreted strictly and does not cover situations where the processing is not genuinely necessary for the performance of a contract, but rather unilaterally imposed on the data subject by the controller." For example, when a controller is contracted to deliver certain goods, it is not possible to use the legal ground of contract in order to carry out profiling, based on the data subject's purchases, as that sort of processing is not necessary for the performance of the contract. Even if these processing activities are mentioned in the fine print it does not make the processing necessary for the performance of the contract.[161] Necessity is therefore "not simply an assessment of what is permitted by or written into the terms of a contract."[162]

Article 6(1)(b) GDPR can be applicable for the processing activities undertaken by companies engaging in the healthcare sector that offer for example electronic devices and health applications, and where the processing of data is necessary for the provision of the devices and apps. In the case of processing data for research purposes, it may be possible for a company to base its processing on this legal basis if the research activities are necessary for the performance of a contract. For example, consider a company that offers an AI-powered health app that provides tailored health advice to its users. To deliver this product, the company must conduct research consisting of the development and training of the algorithms. As the research is necessary for the performance of the contract, it may be possible to base its processing activities on the legal base in Article 6(1)(b) GDPR.

The WP29 has expressed concern over the mixing of legal bases when processing personal data in contractual or quasi-contractual situations and stated that the "two lawful bases for the lawful processing of personal data, i.e. consent and contract cannot be merged and blurred." Thus, when the processing of personal data is necessary for the performance of a contract, consent is not the appropriate legal basis.[163]

[160] EDPB, "Guidelines 2/2019 on the Processing of Personal Data under Article 6(1)(b) GDPR in the Context of the Provision of Online Services to Data Subject" (2019), p. 9.

[161] WP29 (2014), p. 16-17.

[162] EDPB, "Guidelines 2/2019 on the Processing of Personal Data under Article 6(1)(b) GDPR in the Context of the Provision of Online Services to Data Subject" (2019), p. 8.

[163] WP29, "Guidelines on Consent under Regulation 2016/679" (2018), p. 10. This view has later been endorsed by the EDPB, see EDPB, "Guidelines 2/2019 on the Processing of Personal Data under Article 6(1)(b) GDPR in the Context of the Provision of Online Services to Data Subject" (2019), p. 7.

5.3 Task Carried Out in the Public Interest

Another legal ground offered by the GDPR is if the processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller.[164] To be able to invoke this ground, the basis for the processing must be laid down by Union or Member State law.[165] For public universities and higher education institutions, the task to conduct research is set out by the Swedish Higher Education Act, and they can therefore process personal data on the basis of Article 6(1)(e).[166] For other public bodies, it is possible to process data on the same basis, if the task to conduct research is clearly defined by law or a government decision.[167]

For private research entities, the task of conducting research is generally not regulated by law.[168] However, the importance of private researchers being able to invoke the legal basis of public interest has been pointed out by the Swedish government, as research conducted by private entities can still be considered a task of public interest. For example, the preparatory work highlights that higher education institutions can be both public and private and that the task of conducting research carries equal weight. Furthermore, research is often carried out in collaboration between public and private research entities. The Swedish government has also pointed out that it is not always possible or appropriate to obtain consent for the processing of data for research purposes in studies involving a large number of participants. Both public and private research entities therefore need to be able to process personal data without consent.[169]

Since research involving humans is required to carry a theoretical and/or practical value to society at large and because research can only be approved if the risks to the health, safety, and personal integrity of research subjects are outweighed by its scientific value, the government has stated that if a research project has been approved under the Ethical Review Act, it has been deemed to be beneficial to society.[170] Thus, the research is considered a task of public interest and covered by Article 6(1)(e) GDPR.[171] The support in Member State law, which is required when processing data under Article 6(1)(e)

[164] Article 6(1)(e) GDPR.
[165] Article 6(2)(a)-(b) GDPR.
[166] Chapter 1, Section 2 of the Higher Education Act. In Swedish: Högskolelagen (1992:1434).
[167] Prop. 2017/18:298, p. 48.
[168] Ibid., p. 51.
[169] Ibid., p. 50.
[170] Prop. 2002/03:50, p. 98; Section 9 of the Ethical Review Act.
[171] Prop. 2017/18:298, p. 54.

GDPR, is found in the Swedish Act Containing Supplementary Provisions to the EU General Data Protection Regulation.[172]

5.4 Legitimate Interests Pursued by the Controller or by a Third Party

Processing of personal data can be based on the legal ground in Article 6(1)(f) GDPR when processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except when outweighed by the interest or fundamental rights and freedoms of the data subject. The concepts of "interest" and "purpose" are similar, yet distinct. While the purpose of the processing is the specific reason why data is processed, interest is "the broader stake that a controller may have in the processing, or the benefit that the controller derives – or society might derive – from the processing." For example, a company may process personal data for the *purpose* of implementing specific access control procedures, because the company has an *interest* in ensuring the health and safety of its staff.[173]

The legitimate interest must be evident, although not explicitly recognized, in Union or Member State law.[174] Recital 47 GDPR provides two examples of what can constitute legitimate interests which is processing data for the purposes of preventing fraud and for direct marketing purposes. The WP29 has stated that the nature of the interest can vary and may be "compelling and beneficial to society at large" – such as the interest in conducting scientific research when subject to appropriate safeguards – or "less pressing for society as a whole." That might be the case when a company has an economic interest in gaining information about its potential customers to better target advertisements. The WP29 mentions processing for scientific research purposes as one of the most common contexts in which the question of legitimate interest may arise.[175] Since the Swedish government considers the task of conducting research to be of public interest, as mentioned above, the government also assumes that research constitutes a legitimate interest.[176]

[172] According to Chapter 2, Section 2 of the Act Containing Supplementary Provisions to the EU General Data Protection Regulation Personal data can be processed on the basis of Article 6(1)(e) GDPR if the processing is necessary for the performance of a task carried out in the public interest on the basis of a law or regulation, a collective agreement or a decision adopted pursuant to a law or other regulation. For example, a decision under the Ethical Review Act constitutes such a basis.

[173] WP29 (2014), p. 24.

[174] Kotschy (2020), p. 337.

[175] WP29 (2014), p. 24-25.

[176] Prop. 2017/18:298, p. 57.

5.5 Summary and Concluding Remarks

The chapter outlined possible legal bases for companies engaging in data-driven research to base their processing on. One possible legal basis is consent, which must be freely given, specific, informed, and unambiguous. In order for consent to be freely given there cannot be a clear imbalance between the controller and the data subject. According to Recital 43 GDPR, consent is presumed not to be freely given by the data subject when the controller is a public authority. Such a power balance therefore normally does not exist between individuals and companies. Furthermore, consent must be specific, which the EU legislator has recognized can prove difficult in scientific research. Recital 33 therefore allows for data subjects to give consent to "certain areas of scientific research." However, if the processing involves health data, the WP29 has stated that a stricter interpretation of Recital 33 is required. It is not entirely clear where that leaves controllers or whether the flexibility offered by Recital 33 still applies.

Furthermore, it was highlighted that the use of consent as a legal basis may become complicated in big data analytics, as it is not always possible to determine the value that big data holds from the beginning. For example, a company that offers a wearable device in the form of a smartwatch that collects data on sleep, physical activity, heart rate, menstrual cycle, and mental state, might base this collection on the basis of consent, which is given by users upon using the smartwatch. However, as the company explores new ways to make use of this accumulated data, the company may be required to obtain consent again, which can be challenging if not impossible.

Another legal basis examined is that processing is necessary for the performance of a contract to which the data subject is a party. The view of the WP29 and EDPB is that the criteria of necessity must be interpreted strictly and does not cover situations where the processing is not genuinely necessary for the contract. As argued above, this basis may be a possible option for a company that offers AI-powered solutions, such as a health app, that provides health advice to its users and therefore must conduct research consisting of the development and training of algorithms. However, it is not a question of stipulating a number of uses in a contract in order to cover all possible purposes. Given that the use of this legal basis requires that controllers carry out a strict assessment of the necessity requirement, demonstrating that research is necessary for a contract might prove more difficult than rewarding. At the same time the WP29 has stated that when the processing of personal data is necessary for the performance of a contract, consent is not the appropriate legal basis.

The third legal basis examined is if the processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller. In contrast to, for example, public

universities and higher education institutions, the task of conducting research is generally not prescribed by law for private researchers. In Swedish law, research approved under the Ethical Review Act will be covered by the legal basis in Article 6(1)(e) GDPR, as the research task then has been deemed to be in the public interest. This places private researchers on par with public researchers. While this thesis argues that the threshold for research as defined in the Ethical Review Act is higher than that of "scientific research purposes," once researchers have met the definition of research and been approved in an ethical review, they enjoy a certain advantage in that their processing activities are automatically covered by a legal basis in Article 6. The question of legal basis can otherwise be quite difficult to navigate and assess, for example, as shown in relation to the basis concerning contract. Furthermore, researchers engaging in research that involves large datasets can avoid the difficulties connected to the use of consent as a legal basis, such as if a research subject were to withdraw their consent.

The fourth and last legal basis that was examined is processing necessary for the purposes of the legitimate interests pursued by the controller or by a third party. What constitutes a legitimate interest is not defined in the GDPR, however, preventing fraud and direct marketing purposes are provided as examples in Recital 47. The WP29 has stated that a legitimate interest can be more or less pressing for society. In Swedish law, the task of conducting research is assumed to constitute a legitimate interest as it, as explained above, is considered a task of public interest. For companies engaging in data-driven research that meet the definition of research in the Ethical Review Act and have their research approved under an ethical review, the legitimate interest ground will be of less interest, as the processing activities will be covered by the legal basis of a task carried out in the public interest.

Swedish companies engaging in data-driven research, but whose processing activities fall outside the scope of research as defined in the Ethical Review Act, will have to base their processing on another exemption in Article 9(2) GDPR and find an appropriate basis under Article 6(1). This could, for example, also be the case when research falls under the definition in the Ethical Review Act, but research is conducted outside of Sweden and thus, falls outside the territorial scope of the Ethical Review Act. As the processing must be covered by an exemption in Article 9(2) GDPR, obtaining the data subject's consent for the processing is likely to be the primary option. In terms of companies that process vast amounts of personal data, this task may be too costly or even impossible, especially when data has been retrieved from multiple sources.

6 Processing for Scientific Research Purposes: Impact on Other Provisions

This chapter further outlines the regulatory framework for processing for scientific research purposes. As explained above, processing for scientific research purposes provides for a favorable regime in the GDPR. First, the chapter will explain what safeguards apply when processing data for scientific research purposes and the derogations the GDPR allows for in relation to certain rights of the data subject. Secondly, it will outline the exceptions from the purpose and storage limitation principles that are stipulated in the GDPR. Lastly, it will explain the relief from the obligation to provide the data subject with information and the right to erasure. Sweden will be used as a practical example when relevant.

6.1 Safeguards and Derogations

Processing for scientific research purposes under Article 9(2)(j) GDPR must be carried out in accordance with Article 89(1), which requires that safeguards are in place when personal data is processed for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. Article 89(1) further specifies that technical and organizational measures should be implemented, especially to ensure that the principle of data minimization is upheld. Recital 156 clarifies that Member States should provide appropriate safeguards when processing personal data for scientific research purposes. Pseudonymization is explicitly mentioned as an example of a safeguard. Furthermore, the provision refers to "further processing which does not permit or no longer permits the identification of data subjects," which Wiese Svanberg writes could be interpreted as including anonymization. Although pseudonymization and anonymization might be favored, other measures may be appropriate.[177]

As Member States have been given a wide margin of discretion to further specify the purposes of the processing covered by Article 89, for example by further defining what is covered by "scientific research purposes," in practice it is likely that the detailed requirements for the processing will mainly be decided by Member States.[178] In Swedish law, it has been considered sufficient that the GDPR prescribes that personal data must be pseudonymized or subject to other appropriate safeguards when processed for research purposes and it has therefore not been further regulated.[179]

[177] Wiese Svanberg (2020), p. 1247.
[178] Ibid., p. 1246.
[179] Prop. 2017/18:298, p. 70.

The practical value of the safeguards as laid down by Article 89(1) has been questioned and the language of the provision in terms of "safeguards" has been criticized as being "short and vague."[180] As will be developed below, processing for scientific research purposes allows for some relief in the application of the principles concerning purpose and storage limitations. Although the GDPR admits relief from these principles in relation to scientific research, the processing of data still has to adhere to the safeguards under Article 89(1). Pormeister argues that since there is no guarantee of appropriate safeguards, as their implementation is left to the discretion of the Member States, the relief from the principles of purpose and storage limitations would still apply as they are laid down in directly applicable provisions, even if Member States have failed to implement safeguards.[181]

In contrast to the requirements of "appropriate safeguards," Article 89 also allows Member States to provide for derogations from the rights referred to in Article 15 (right of access by the data subject), Article 16 (right to rectification), Article 18 (right to restriction of processing), and Article 21 (right to object), when processing data for scientific or historical research purposes or statistical purposes "in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes." As the implementation of derogations is left to the discretion of Member States, Pormeister states that this may create "a forum-shopping syndrome," where data processing activities are carried out in the Member States which has implemented the most derogations.[182]

In Swedish law, a general exemption from the right in Article 15 GDPR exists in the Act Containing Supplementary Provisions to the EU General Data Protection Regulation, which stipulates that a data subject's right to access the personal data does not apply when the controller is prohibited, for example by law or other regulation, to disclose the data.[183] For example, information about an individual's health status can in certain cases be kept secret from them.[184] The provision further states that if the controller is not an authority, the exemption also applies to information that would have been confidential under the Swedish Public Access to Information and Secrecy Act. However, no further derogation beyond this general exemption for scientific research purposes has been implemented in Swedish law.[185]

Neither has a derogation from the right to rectification in Article 16 GDPR been introduced. According to Article 16, the data subject has the right to obtain from the controller without undue delay the rectification of inaccurate

[180] Pormeister (2017), p. 140.
[181] Pormeister (2017), p. 139-140.
[182] Ibid., p. 140.
[183] Chapter 5, Section 1 of the Act Containing Supplementary Provisions to the EU General Data Protection Regulation; see also Article 23 GDPR.
[184] See Chapter 25, Section 6 of the Public Access to Information and Secrecy Act.
[185] Prop. 2017/18:298, p. 119.

personal data concerning him or her and has the right to have incomplete personal data completed. The preparatory work states that although this right may affect the conduct of research, it will not be affected to such an extent that the conditions for introducing a derogation are fulfilled.[186] No derogation from the right to restriction of processing has been introduced either.[187]

Regarding the possibility to introduce a derogation from the right to object to the processing of personal data in Article 21 GDPR, the preparatory work points out that such a derogation already exists in Article 21(6), which stipulates that the right to object the processing for scientific or historical research purposes or statistical purposes does not apply if the processing is necessary for the performance of a task carried out for reasons of public interest. As outlined in section 5.3, research approved under the Ethical Review Act is considered a task of public interest. It has therefore not been considered necessary to impose further restrictions to this right.[188]

6.2 The Purpose Limitation Principle

The purpose limitation principle is laid down in Article 5(1)(b) GDPR. It consists of two components: Personal data must be collected for specified, explicit, and legitimate purposes ("purpose specification") and not further processed in a way that is incompatible with the initial purposes ("compatible use"). When further processing of personal data is not based on the consent of the data subject or Union or Member State law, the controller must assess whether the new purpose is compatible with the purpose for which the data was collected. To decide this, the controller should take into account, inter alia, if there is any link between the new and the original purposes, the context in which the personal data was collected, the nature of the personal data, in particular if it involves sensitive data such as health data, the possible consequences of the intended further processing for the data subject, and the existence of appropriate safeguards.[189] Thus, the provision seeks to bring some predictability for the data subject while also leaving room for flexibility for the re-use of data by the controller.[190] Where further processing is considered incompatible with the original purpose and EU or Member State law does not contain a specific provision allowing the incompatible further processing, the controller must obtain the data subject's consent in order to pursue the additional purpose.[191]

The provision on purpose limitation contains a legal presumption that further processing for archiving purposes in the public interest, scientific or historical

[186] Ibid., p. 123.
[187] Ibid., p. 127.
[188] Ibid., p. 129-130.
[189] Article 6(4) GDPR.
[190] Article 5(1)(b); Ivanova (2019), p. 109.
[191] Kotschy (2020), p. 344.

research purposes or statistical purposes, if it is carried out in accordance with Article 89(1), shall not be considered to be incompatible with the initial purposes.[192] Further processing for scientific research purposes is therefore particularly favored under the GDPR.[193]

6.2.1 Purpose Specification

The principle of purpose limitation is a cornerstone in EU data protection law and is also explicitly stated in Article 8(2) of the EU Charter of Fundamental Rights. When individuals entrust others with their data, they usually do so with an expectation of the purposes for which the data will be used. Meeting these expectations is important in maintaining trust and legal certainty.[194] However, the principle can be said to be in direct conflict with big data analytics, which has been discussed by several scholars.[195] One challenge especially with big data analytics or research is to specify the purpose in a sufficient way, as the processing of big data can have an almost infinite number of uses that cannot be identified and articulated to the data subject at the time of data collection.[196] The purpose may also change as the machine learns and develops.[197] Furthermore, the very value with big data analytics can be said to be its ability to process data for different purposes and to analyze data in ways that may have not been envisaged at the time of data collection, making the principle of purpose limitation an obstacle to extracting its value.[198]

When further processing of personal data is not considered compatible with the initial purposes, entities engaging in big data analytics will have to carefully monitor their practices to ensure that they do not process data outside of the specified purposes. This may prove both costly and difficult – if not even impossible, Zarsky argues. Furthermore, it would most likely not be possible to circumvent the rule by defining the purpose in a too vague or general manner, as the purpose must be "specific."[199]

Zarsky states that the purpose principle can, at least on a theoretical level, ensure some control for the data subject that their data is used for the intended purposes. Furthermore, it can foster trust in the data environment as well as encourage competition, as it weakens the position of monopolies in the data market and allows startups to enter and compete. However, it might also hinder competition, as it may constitute an obstacle for startups to obtain

[192] Article 5(1)(b) GDPR.
[193] Prop. 2017/18:298, p. 60.
[194] WP29 (2013), p. 4.
[195] See e.g. Ivanova, Re-using Personal Data for Statistical and Research Purposes (2019); Pierce, Machine Learning for Diagnosis and Treatment: Gymnastics for the GDPR (2018); Zarsky, Incompatible: The GDPR in the Age of Big Data (2017).
[196] Pierce (2018), p. 339.
[197] Ivanova (2019), p. 111.
[198] Ibid.; Zarsky (2017), p. 1005-1006.
[199] Zarsky (2017), p. 1006.

personal data on secondary markets. Zarsky concludes that even taking into account the flexibility that is offered when further processing is compatible with the initial purposes, the GDPR clearly creates an obstacle for big data analytics.[200]

6.2.2 Compatible Use

As described above, the provision on purpose limitation contains a legal presumption that further processing for scientific research purposes, in accordance with Article 89(1), is considered to be compatible with the initial purpose.[201] In that case, Recital 50 states that "no legal basis separate from that which allowed the collection of the personal data in the first place is required."

A large part of healthcare data will be generated from various devices that collect a substantial amount of information about their users. Today, almost everyone has a smart phone and/or other smart device that registers data that can be useful from a health perspective, for example information about mobility and sleep patterns. There is also an increasing number of devices for health and fitness, such as fitness bands, that produce large amounts of data over time. Furthermore, mobile apps for different medical conditions constitute an important new source of data for self-managing one's health.[202] The favored position of scientific research purposes in the GDPR allows for companies within the health sector to further process their accumulated data for scientific research purposes, without requiring a separate legal basis from that which allowed the collection in the first place.

It is less clear whether a controller who collects personal data from another controller for scientific research purposes needs a separate legal basis under Article 6(1) GDPR for the collection and further processing, or if it can rely on the legal basis for the initial collection.[203] The Swedish government has stated that the transfer of data from one controller to another in order for the latter to process the data for scientific research purposes constitutes further processing. However, obtaining personal data for scientific research purposes does not constitute further processing but instead a collection of personal data. Thus, for the collection of personal data to be lawful, the controller receiving the data must be able to invoke its own legal basis under Article 6(1) GDPR.[204] Furthermore, in the case of processing of special categories of

[200] Zarsky (2017), p. 1006-1007.
[201] Article 5(1)(b) GDPR.
[202] Bohr and Memarzadeh (2020), p. 14-15.
[203] Öman (2023), legal commentary on Article 5(1)(b) GDPR.
[204] Prop. 2017/18:298, p. 61.

personal data, such processing is generally prohibited and must therefore always be covered by one of the exemptions under Article 9(2) GDPR.[205]

The view of the Swedish government is not generally agreed upon and the EDPB has been requested by the European Commission to clarify to what extent the initial legal basis can be relied upon for the further processing in cases where data is re-used in different research projects of the same nature.[206] Öman states that the wording of Recital 50 and "the systematics of the GDPR" do not exclude the possibility of using the legal basis of the original controller when data is re-used by a different controller, but that this may not apply when sensitive data is collected from another controller.[207] The EDPS has stated that "in principle personal data collected in the commercial or healthcare context, for example, may be further used for scientific research purposes, by the original or a new controller, if appropriate safeguards are in place," which indicates that it interprets the re-use of data by the new controller as further processing.[208]

Becker et al. write that different interpretations are possible given how you view the data lifecycle. With a "data-focused view," the first collection (i.e. directly from the data subject) for specific purposes is seen as "initial," which implies that any subsequent processing for different purposes constitutes further processing under the GDPR, irrespective of the controller. This view can be supported by the wording in Recital 50: "processing of personal data for purposes other than those for which the personal data were initially collected."[209] In contrast, with a "controller-focused view," a data lifecycle begins when a controller collects personal data – either directly from the data subject or from another source (such as from a different controller) – and ends with the fulfillment of the purpose(s) for which the controller collected the data.[210] Therefore, each time a controller collects data, whether directly from the data subject or from another controller, it marks the start of the primary processing.[211]

Becker et al. argue that further processing in the GDPR "is to be understood in relation to the purpose for which a particular controller originally collected the data, whether directly from the data subject or by obtaining existing data from another source." This interpretation is in line, they argue, with the wording in Article 5(1)(b) GDPR, which they mean refers to "data collection in

[205] Öman (2023), legal commentary on Article 5(b) GDPR; SOU 2017:50, p. 147.

[206] EDPB (2021), p. 6. The EDPB answered that it was going to provide more clarification regarding this matter in its Guidelines on processing personal data for scientific research purposes. As of April 2024, these guidelines have not yet been published.

[207] Öman (2023), legal commentary on Article 5(b) GDPR.

[208] EDPS (2020), p. 22.

[209] Becker et al. (2023), p. 152. However, the authors argue that the wording "initially collected" in Recital 50 is "easily misleading."

[210] Ibid., p. 138.

[211] Ibid., p. 151.

general, as opposed to data collection directly from the data subject."[212] The term "initial purposes" in Article 5(1)(b) therefore refers to any collection of data for a specific purpose, whether the collection of data is directly from the data subject or if existing data has been obtained from another source.[213] Furthermore, Becker et al. argue that, as the recitals are non-binding, it does not override the requirement of a legal basis under Article 6 GDPR, even if the further processing is done by the original controller. When a controller further processes data for scientific research purposes it must therefore always be based on a legal ground under Article 6.[214]

Even though different interpretations are possible and exist in the literature, the Swedish government's interpretation constitutes an important limitation on the sharing of health data between different controllers for scientific research purposes when research is conducted in Sweden.

6.3 The Storage Limitation Principle

According to the storage limitation principle in Article 5(1)(e) GDPR, personal data must not be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data is processed. Recital 39 further states that time limits should be established by the controller for erasure or for a periodic review, in order to ensure that that personal data is not kept longer than necessary. Processing of personal data solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes is exempted from the storage limitation principle, as long as it is carried out "in accordance with Article 89(1) subject to implementation of the appropriate technical and organizational measures required by the GDPR in order to safeguard the rights and freedoms of the data subject."[215]

For certain machine learning programs that perform best with the maximum amount of personal data, "the value of data in this identifiable form is precisely what makes it valuable." The more information that is available about, for example, each patient, the more accurate the predictions or decisions the programs will be able to present. However, storing personal data for longer time periods than necessary for the treatment of a patient, for example, may be in violation of the storage limitation principle.[216]

While the GDPR provides an exemption from the storage limitation principle in relation to scientific research, the principle may have an impact on the ability to develop machine learning programs that rely on vast amounts of

[212] Ibid., p. 138.
[213] Ibid., p. 151.
[214] Ibid., p. 149.
[215] Article 5(1)(e) GDPR.
[216] Pierce (2018), p. 340.

personal data, if data is removed as soon as the processing for purposes other than scientific research have been achieved.[217]

6.4 Obligation to Provide the Data Subject with Information

The processing of personal data for scientific research purposes affects the application of Article 14 GDPR, which sets forth an obligation for the controller to provide the data subject with information, such as the purpose of the processing and the source from which the data originates, when personal data has not been obtained directly from the data subject. This may be the case, for example, when a research entity retrieves personal data from another party, such as a company or a healthcare provider. According to Article 14(5)(b), the obligation to provide the data subject with information does not apply if the provision of such information proves impossible or would involve a disproportionate effort, in particular for processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, subject to conditions and safeguards under Article 89(1) or in so far as the obligation is likely to render impossible or seriously impair the achievement of the objectives of that processing.

To determine if the provision of information would involve a disproportionate effort, the number of data subject should, inter alia, be taken into consideration.[218] In the context of data-driven research, the requirement to provide data subjects with information may render a disproportionate effort (if not impossible) due to the large volume of data, which may include data from not only an enormous number of subjects but also from many different sources. The WP29 has stated that "the exception cannot be routinely relied upon by data controller who are not processing personal data for the purposes of archiving in the public interest, for scientific or historical research purposes or statistical purposes," indicating that entities processing personal data for scientific research purposes may more successfully invoke the exemption from the obligation to provide information.[219]

As further processing for scientific research purposes is deemed compatible with the initial purpose, collected data could essentially be further processed for scientific research without the data subject's knowledge, if Article 14(5)(b) is applicable.[220] Taking into account the interpretation of the purpose limitation principle in Swedish law, this may be the case concerning research in Sweden, when collected data is transferred to another controller for processing for scientific research purposes. However, in the preparatory work,

[217] Pierce (2018), p. 341.
[218] Recital 62 GDPR.
[219] WP29, "Guidelines on Transparency under Regulation 2016/679," (2018), p. 27
[220] Pormeister (2017), p. 140.

the government has made the assessment that an ethical review board has the option, in connection with an approval, to impose conditions that information must be provided to the data subject if this is considered an appropriate safeguard under Article 89(1) GDPR.[221]

6.5 Right to Erasure ("Right to be Forgotten")

The right to erasure or "right to be forgotten" in Article 17, establishes a right for the data subject to have his or her personal data erased without undue delay under certain conditions. This could be the case when the personal data is no longer necessary in relation to the purposes for which it was collected or otherwise processed or if the data subject withdraws consent on which the processing is based according to Article 6(1)(a) or 9(2)(a), and where there is no other legal ground for the processing.[222] Furthermore, if the controller has made the personal data public and is obliged to erase the personal data, the controller must, with regard to available technology and the cost of implementation, inform other controllers which are processing the personal data about the request of the data subject to erase the data.[223]

When personal data is processed for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1), Article 17(3)(d) GDPR provides for an exemption from the right to erasure, in so far as the right is likely to render impossible or seriously impair the achievement of the objectives of that processing. Pormeister states that while it may be difficult or even impossible to retroactively erase personal data when scientific research has been conducted, it might be possible to successfully claim the erasure of personal data when it is kept for future research.[224]

6.6 Summary and Concluding Remarks

The chapter further outlined the regulatory framework for processing for scientific research purposes. This section aims to answer the second research question: How do the GDPR's scientific research regime and its implementation in Swedish law balance the interests of data subjects against the interests of controllers, and how might this balance affect data-driven research? It should be mentioned again that the thesis is primarily interested in which ways the scientific research regime allows for companies to undertake activities that would otherwise conflict with the interests of data subjects.

[221] Prop. 2017/18:298, p. 115.
[222] Article 17(1)(a)-(b) GDPR.
[223] Article 17(2) GDPR.
[224] Pormeister (2017), p. 140.

Firstly, the chapter outlined Article 89 GDPR, which requires safeguards and allows for derogations relating to the processing for scientific research purposes. According to the provision, safeguards in the form of technological and organizational measures, such as pseudonymization, must be in place when processing data for scientific research purposes. Recital 156 further states that Member States should provide for those measures. Pormeister argues that because the implementation of the safeguards are left to the Member States, the requirement of safeguards in the GDPR may not have an effect in practice.

In Swedish law, the requirement of safeguards as stipulated in Article 89(1) GDPR has been found enough, and no specific provision in terms of safeguards have been implemented in Swedish law. While that could suggest a risk for the data subject, in terms of research involving sensitive data, ethical review would work as a safeguard to ensure that controllers have technical and organizational measures in place. The effect suggested by Pormeister would likely not play out for research conducted in Sweden, although the lack of specific implementation do leave some room for controllers in deciding what measures might be appropriate. The requirement of ethical review ensures that there is a balance between the two interests, or at least, that there is a scientific value that motivates the risk to the personal integrity However, ethical review affects the flexibility offered by the scientific research regime as companies that want to conduct research must submit an application for ethical review and cannot proceed until the project has been approved.

Article 89(2) further allows for derogations from the right of access, right to rectification, right to restriction, and right to object, when processing data for scientific research purposes, thereby granting Member States discretion to balance the data subject's interest in maintaining these rights and the controller's interest in conducting research. While allowing for derogations from data subjects' rights may appear to strongly favor the controller, the GDPR stipulates a high threshold for this to apply: Derogations are only permissible in so far as the rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfillment of those purposes. Sweden has chosen not to enact any derogations under Article 89(2), deciding in advance that it does not affect processing for research to the extent that the conditions in Article 89(2) are met. For example, regarding the right to rectification, the preparatory work states that although this right may affect the conduct of research, it does not affect it to the extent that a derogation should be introduced.

As every research project is different, it is difficult to see how it is possible to assess in advance whether the rights affect the possibility to carry out research. Another option, which would better balance the rights of data subjects

and controllers, would have been to introduce in Swedish law the possibility of derogating from the rights together with the requirement in Article 89(2), i.e. that the right must render impossible or seriously impair the research, thereby deciding on a case-by-case basis the effect that the data subjects' rights may have on the conduct of research. Using the example of the right to rectification, while there is at the same time an interest for the controller in ensuring that the data is correct, the exercise of this right may affect data-driven research where results are based on the integration of data from multiple sources. As Swedish law does not provide for the possibility of derogating from this right, controllers must however comply with it, for example by ensuring that there are mechanisms in place in case of such requests.

Secondly, the chapter examined the purpose limitation principle in Article 5(1)(b) GDPR, which states that data must be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes. The purpose limitation principle clearly favors the data subject, ensuring that individuals can trust that their data is not processed in ways that go beyond their reasonable expectations. As stated in section 6.2.1, the purpose limitation principle is in conflict with big data analytics, which may adversely affect companies engaging in data-driven research. Not only does the purpose limitation principle pose challenges in defining the purpose, as big data analytics can have an almost enormous amount of uses, but it might also create an obstacle as part of its value lies in its ability to analyze data for different purposes and in ways that have not been obvious from the outset.

While the purpose limitation principle primarily creates an obstacle for companies engaging in data-driven research, it might at the same time have a positive impact on competition. As Zarsky interestingly points out, the principle may weaken the position of monopolies in the data market and allow for startups to enter. As illustrated by the examples of the Google DeepMind patient data deal and the genetic testing company, presented in section 1.1, there is a high demand and great economic value associated with health data. The purpose limitation principle can therefore allow for fair competition by ensuring that not only the largest actors with the most resources have access to health data. However, while the issue of large corporations monopolizing the market should not be undermined, the purpose limitation principle is also a crucial obstacle for smaller companies wanting to engage in research that benefits from vast amounts of data, as it might be difficult to outline from the outset what data is needed and defining the purposes in a manner that is not too general. This barrier is likely more significant than the issue of access to health data, especially as there has been a surge in health data over the last decade, which was also highlighted in the introduction to the thesis.

Furthermore, the second part of the purpose limitation principle was examined, namely that data cannot be further processed in a manner that is

incompatible with the initial purposes. The provision in Article 5(1)(b) GDPR contains a legal presumption that data processed for scientific research purposes is compatible with the initial purposes. According to Recital 50, when "processing is considered compatible with the purposes for which the personal data were initially collected […] no legal basis separate from that which allowed the collection of the personal data is required."

The legal presumption in regard to scientific research purposes appears at first sight to shift the balance of interests significantly towards the controller. However, the extent of the relief from this principle is not entirely clear. What creates the confusion is the wording of Recital 50 which speaks of data that has been "initially collected" instead of "collected." While the general opinion appears to be that the provision of data from one controller to another is considered further processing and does not require a separate basis from the initial one, there are different interpretations as to whether the new controller that obtains data can also rely on the legal basis of the original controller.

The EDPS has stated, again in a preliminary opinion, that "in principle personal data collected in the commercial or healthcare context, for example, may be further used for scientific research purposes, by the original or a new controller, if appropriate safeguards are in place." While this statement suggests that the legal presumption applies to both the original and new controller, it does not settle the debate. What furthermore contributes to the confusion, apart from the wording of Recital 50, is that the legal presumption is essentially an exemption from the purpose limitation principle, whereas Recital 50 extends this to also concern the requirement of a legal basis in Article 6. The statement by the EDPS could therefore be interpreted as meaning that while the processing by the new controller is considered further processing, i.e. the new controller does not have to carry out an assessment of whether the original and new purposes are compatible, it does not necessarily mean that the second controller does not need to invoke its own legal basis.

As mentioned above, Becker et al. have examined the question of whether processing by the new controller is considered further processing from two different perspectives, a data-focused view and a controller-focused view, while favoring the latter. With a data-focused view, the first collection of data is seen as the initial, and every subsequent processing for scientific research purposes is seen as compatible further processing. With a controller-focused view, each time a controller collects data, whether directly from the data subject or from another controller, it marks the start of the primary processing. In Swedish law, the controller-focused view has been adopted, which means that the collection of data from another controller for scientific research purposes does not constitute further processing, but instead marks the start of primary processing.

While Becker et al. are correct that the requirement of legal basis in Article 6 cannot be overridden by the recitals, there is not necessary a conflict between Article 6 and Recital 50. As long as the purposes of the collection and further processing are compatible, further processing is not exempted from the requirement of legal basis, but covered by the legal basis pertaining to the original processing. Thus, when further processing is not considered compatible with the initial purposes the controller must establish a legal basis for the processing for this new purpose.

Requiring a controller to establish a new legal basis for every different but compatible purpose would also be quite burdensome. Recital 50 can be interpreted as implying that if the different but compatible purpose was defined from the beginning, the data subject would have consented to it or it would have been covered by the legal basis relating to legitimate interests, for example. Thus, taking into account the interest of the data subject in not having its data processed for purposes that go beyond their reasonable expectations, it is fair for controllers not having to establish a new legal basis for further processing, which is also in accordance with the language of Recital 50.

In light of the privileged position of scientific research in the GDPR and the language of Recital 50, it is unclear whether the Swedish interpretation aligns with the intentions of the EU legislator. However, the Swedish interpretation may not have far reaching consequences for a controller that obtains health data for scientific research purposes, regardless of if it constitutes further processing or a collection of data, since the controller must obtain approval through an ethical review before proceeding. Once research has been approved through an ethical review it is automatically covered by the legal basis of a task carried out in the public interest. The question of whether the legal basis of the original controller applies for the processing of the second controller or whether a new legal basis has to be invoked therefore becomes irrelevant. It could have an effect on processing for scientific research purposes that does not involve special categories of health data, however, that falls outside the scope of this thesis. The flow of personal data for scientific research purposes that the relief from the purpose limitation principle allows for is hindered, however, because of the requirement of ethical review in Swedish law, as a controller cannot go forward with the collection or other processing of the data without obtaining a review.

Thirdly, the thesis outlined the storage limitation principle, which, first and foremost favors the data subject's interest in not having its data stored in a form which permits identification for longer periods than is necessary. However, the provision provides an exemption when data is processed solely for scientific research purposes. There are some difficulties regarding how this provision should apply. It is unclear, for example, whether data can be stored for up to one year or even several years. Furthermore, there is again a conflict between processing that involves a substantial amount of personal data and

the provisions in the GDPR. As the exception only applies to data which will be processed *solely* for scientific research purposes, it might be difficult to decide in advance what data will be useful for future research.

For example, consider a company that offers different types of consumer technology products, which collect a significant amount of personal data, including health data. The company may store this data for as long as the consumer uses the products, such as a mobile app. While this data could be useful from a research perspective, it is challenging to decide which data will be used solely for scientific research purposes, as some data may not be processed for research purposes at all and some data might be valuable for research but also processed in order to enhance a tool's user-friendliness or improve existing features, without classifying as scientific research. While the relief from the storage principle shifts the balance of interests towards the controller, the criterion of "solely" makes it a quite narrow exception, which might affect research practices that benefit from large amounts of data. This issue might be more evidently for companies engaging in commercial activities, where data is primarily collected for a different purpose than research.

Furthermore, Article 14 GDPR was examined. Companies engaging in data-driven research that involves data from a large number of subjects and sources are likely to be able to successfully invoke the exemption from the obligation to provide the data subject with information in Article 14 GDPR. While the provision allows for an exception to a data subject right, the provision implements a more practical perspective rather than simply favoring the controller's interest. For research that is carried out in Sweden, whether controllers are exempted from this obligation or not may be left in the hands of the Ethical Review Authority, as the preparatory work highlights the possibility of imposing conditions that information must be provided to the data subject along with an approval.

Finally, the chapter outlined the right to erasure or the "right to be forgotten," which allows for the data subject to have his or her data immediately deleted upon request. Processing for scientific research purposes allows for an exception to this right, which again, although the exemption favors the interests of the controller of conducting research, contains a high threshold for this to apply: the right must render impossible or seriously impair the achievement of the research. Using the example of machine learning programs that perform best with the maximum amount of personal data and is able to present more accurate predictions and decisions the more information that is available, companies engaging in data-driven research may be successful in claiming that the right to be forgotten affects the research to the extent that the conditions in Article 17(3)(d) GDPR are met. However, it is a narrow exemption, and it might be difficult for companies to argue that the conditions are met.

7 Conclusions

This chapter aims to bring together the conclusions that have been made in regard to the research questions and offers some reflections in relation to the findings of the thesis. The first research question was: To what extent can Swedish companies claim the scientific research exemption in Article 9(2)(j) GDPR when processing health data? As concluded in Chapters 3 and 4, neither the GDPR nor the Ethical Review Act preclude private entities or activities that are undertaken with a commercial interest. The focus of the definition of research in the Ethical Review Act, of acquiring new knowledge and the theoretical and/or practical value of the research, sets forth a higher threshold of what constitutes research than the GDPR. Companies that are primarily driven by commercial interests might have difficulties in clarifying the scientific value of their activities. For example, what might fall under technological development in the GDPR might not be defined as research under the Ethical Review Act without a "significant improvement to an existing system."

A clear drawback with ethical review as a specific and suitable measure required to benefit from the scientific research exemption, is that the definition of research in the Ethical Review Act and its territorial scope do not align with that of the GDPR. This discrepancy can present challenges for determining what activities benefit from the exemption, particularly in international research projects where part of the research is conducted in Sweden. This creates a fragmented legal framework in regard to the scientific research regime within the EU.

It is debatable whether ethical review under the Ethical Review Act is always an appropriate measure. While this thesis does not seek to critique the ethical review system per se, the requirement that all research involving health data be subject to ethical review may, as health data encompasses a wide range of types of information, result in research with a low risk of harm being subject to ethical review. This may be burdensome in terms of time and resources, particularly for startups, which might refrain from undertaking their activities. For example, information about people's weight, blood pressure, and physical activity might involve a low risk of harm, but still hold considerable value for the development of tools for health monitoring. As the data, taken together, would probably constitute health data within the meaning of the GDPR, a company that wants to process this data for research automatically has to undergo an ethical review in order to benefit from the scientific research exemption. The need for an ethical review may be more apparent regarding information about whether a person carries a sexually transmitted disease or information about a person's mental health status, as it might involve a higher risk of harm if revealed. However, determining what health data requires ethical review or not would also pose significant challenges.

A further complicating factor for companies, especially for companies which primarily collect health data for a purpose other than research, is that they cannot proceed to use this data immediately, but must first have their research approved through an ethical review. Moreover, if, after obtaining ethical approval, the company then wishes to re-use this data but for another research project, it must obtain ethical approval again. Arguably, this removes some of the flexibility that the scientific research exemption otherwise intends to provide for scientific research.

The second research question was: How do the scientific research regime in the GDPR and its implementation in Swedish law balance the interests of data subjects against the interests of controllers, and how might this balance affect data-driven research? As outlined in Chapter 6, the scientific research regime appears at first sight to shift the balance of interest significantly in favor of the controller. However, for this to apply, it is often required that the application of the provisions would render impossible or seriously impair the achievement of scientific research, thereby narrowing the scope of the framework. Sweden has also adopted a rather passive stance in terms of legislation, for example by refraining from introducing the possibility to derogate from certain rights of the data subject. While this may adversely affect the flexibility of companies engaging in data-driven research, the biggest challenge is evidently to overcome the conflict between the provisions of the GDPR and data-driven research involving substantial amounts of personal data.

Bibliography

Books

Bentzen, Heidi Beate. “In the Name of Scientific Advancement: How to Assess What Constitutes ‘Scientific Research’ in the GDPR to Protect Data Subjects and Democracy.” In *Disinformation and Digital Media as a Challenge for Democracy*, edited by Georgios Terzis et al. Intersentia, 2020.

Bohr, Adam & Memarzadeh, Kaveh. “Current Healthcare, Big Data, and Machine Learning.” In *Artificial Intelligence in Healthcare*, edited by Adam Bohr & Kaveh Memarzadeh. Academic Press, 2020.

Bohr, Adam & Memarzadeh, Kaveh. “The Rise of Artificial Intelligence in Healthcare Applications.” In *Artificial Intelligence in Healthcare*, edited by Adam Bohr & Kaveh Memarzadeh. Academic Press, 2020.

Bygrave, Luca & Tosoni, Lee A. “Article 4(1). Personal Data.” In *The EU General Data Protection Regulation (GDPR): A Commentary*, edited by Christopher Kuner et al. Oxford University Press, 2020.

Bygrave, Luca & Tosoni, Lee A. “Article 4(2). Processing.” In *The EU General Data Protection Regulation (GDPR): A Commentary*, edited by Christopher Kuner et al. Oxford University Press, 2020.

Bygrave, Luca & Tosoni, Lee A. “Article 4(7). Controller.” In *The EU General Data Protection Regulation (GDPR): A Commentary*, edited by Christopher Kuner et al. Oxford University Press, 2020.

Bygrave, Luca & Tosoni, Lee A. “Article 4(8). Processor.” In *The EU General Data Protection Regulation (GDPR): A Commentary*, edited by Christopher Kuner et al. Oxford University Press, 2020.

Bygrave, Luca & Tosoni, Lee A. “Article 4(15). Data Concerning Health.” In *The EU General Data Protection Regulation (GDPR): A Commentary*, edited by Christopher Kuner et al. Oxford University Press, 2020.

Georgieva, Ludmila & Kuner, Christopher. “Article 9 Processing of Special Categories of Personal Data.” In *The EU General Data Protection Regulation (GDPR): A Commentary*, edited by Christopher Kuner et al. Oxford University Press, 2020.

Gerke, Sara. et al. “Ethical and Legal Challenges of Artificial Intelligence-driven Healthcare.” In *Artificial Intelligence in Healthcare*, edited by Adam Bohr & Kaveh Memarzadeh. Academic Press, 2020.

Hlávka, Jakub P. “Security, Privacy, and Information-Sharing Aspects of Healthcare Artificial Intelligence.” In *Artificial Intelligence in Healthcare*, edited by Adam Bohr & Kaveh Memarzadeh. Academic Press, 2020.

Ivanova, Yordanka. “Re-using Personal Data for Statistical and Research Purposes.” In *Privacy Technologies and Policy*, edited by Maurizio Naldi et al. Springer, 2019.

Kosta, E. “Article 7 Conditions for Consent.” In *The EU General Data Protection Regulation (GDPR): A Commentary*, edited by Christopher Kuner et al. Oxford University Press, 2020.

Kotschy, W. “Article 6 Lawfulness of Processing.” In *The EU General Data Protection Regulation (GDPR): A Commentary*, edited by Christopher Kuner et al. Oxford University Press, 2020.

Krzysztofek, Mariusz. *GDPR Personal Data Protection in the European Union*. Wolters Kluwer, 2021.

Meszaros, Janos et al. “Nudging Consent and the New Opt-Out System to the Processing of Health Data in England.” In *Legal Tech and the New Sharing Economy*, edited by Marcelo Corrales Compagnucci. Springer, 2020.

Riesenhuber, Karl. “Interpretation of EU Secondary Law.” In *European Legal Methodology*, edited by Karl Riesenhuber. Intersentia, 2nd ed., 2021.

Slokenberga, Santa. “Setting the Foundations: Individuals Rights, Public Interest, Scientific Research and Biobanking” in *GDPR and Biobanking*, edited by Santa Slokenberga et al. Springer, 2021.

Verhenneman, Griet. *The Patient, Data Protection and Changing Healthcare Models: The Impact of e-Health on Informed Consent, Anonymisation and Purpose Limitation*. Intersentia, 2021.

Wiese Svanberg, Christian. “Article 89 Safeguards and Derogations Relating to Processing for Archiving Purposes in the Public Interest, Scientific or Historical Research Purposes or Statistical Purposes.” In *The EU General Data Protection Regulation (GDPR): A Commentary*, edited by Christopher Kuner et al. Oxford University Press, 2020.

Öman, Sören. *Dataskyddsförordningen (GDPR) m.m. En kommentar*. Nordstedts Juridik, 2nd ed., 2023.

Academic Journals

Becker, R. et al., "Secondary Use of Personal Health Data. When is it 'Further Processing' under the GDPR and What are the Implications for Data Controllers," 30(2) *European Journal of Health Law* (2023).

Cate, F. H. & Mayer-Schönberger, V., "Notice and Consent in a World of Big Data," 3(2) *International Data Privacy Law* (2013).

Comandè, G. & Schneider, G., "Differential Data Protection Regimes in Data-Driven Research: Why the GDPR is More Research-Friendly Than You Think," 23 *German Law Journal* 559 (2022).

Determann, L., "Healthy Data Protection," *Michigan Technology Law Review* (2020).

Ducato, R., "Data Protection, Scientific Research, and the Role of Information," *Computer Law & Security Review* (2020).

Holtz, H. & Ledendal, J., "Överlappningen mellan dataskydd och marknadsrätt Dataskyddsförordningens tillämpning på marknadsföring och marknadsrättens tillämpning på kommersiell personuppgiftsbehandling," *SvJT* 2020 s. 140.

Kruus, M., "The Public Interest Requirement in the Secondary Use of Health Data in Scientific Research: The Examples of Estonia and Finland," *Juridica International* 32 (2023).

Liangyuan, N. et al., "Feasibility of Reidentifying Individuals in Large National Physical Activity Data Sets From Which Protected Health Information Has Been Removed With Use of Machine Learning," 1(8) *JAMA Netw Open* (2018).

Mészáros, J. & Ho, C., "AI Research and Data Protection: Can the Same Rules Apply for Commercial and Academic Research under the GDPR," 41 *Computer Law & Security Review* (2021).

Mészáros, J. & Ho, C., "Big Data and Scientific Research: The Secondary Use of Personal Data under the Research Exemption in the GDPR" (4) *Hungarian Journal of Legal Studies* (2018).

Molnár-Gábor et al., "Harmonization After the GDPR? Divergences in the Rules for Genetic and Health Data Sharing in Four Member States and Ways to Overcome them by EU Measures: Insights from Germany, Greece, Latvia and Sweden," 84 *Seminars in Cancer Biology* (2022).

Nicholson Price, W. & Glenn Cohen, I., "Privacy in the Age of Medical Big Data," 25 *Nature Medicine* (2019).

Ohm, P., "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization," in 57 *UCLA Law Review* (2010).

Pormeister, K., "Genetic Data and the Research Exemption: Is the GDPR Going Too Far," 7(2) *International Data Privacy Law* (2017).

Pierce, R., "Machine Learning for Diagnosis and Treatment: Gymnastics for the GDPR," 4(3) *European Data Protection Law Review* (2018).

Quinn, P. & Quinn, L., "Big Genetic Data and its Big Data Protection Challenges," 34(5) *Computer Law & Security Review* (2018).

Wouters et al., "Putting the GDPR into Practice: Difficulties and Uncertainties Experienced in the Conduct of Big Data Health Research," 7(2) *European Data Protection Law Review* (2021).

Zarsky, T., "Incompatible: The GDPR in the Age of Big Data" 47(4) *Seton Hall Law Review* (2017).

Papers of Data Protection Authorities

EDPB 2018: European Data Protection Board "Endorsement 1/2018" (2018).

EDPB 2019: European Data Protection Board "Guidelines 2/2019 on the Processing of Personal Data under Article 6(1)(b) GDPR in the Context of the Provision of Online Services to Data Subject" (2019).

EDPB 2019: European Data Protection Board "Guidelines 3/2018 on the Territorial Scope of the GDPR (Article 3)" (2019).

EDPB 2020: European Data Protection Board "Guidelines 05/2020 on Consent under Regulation 2016/679" (2020).

EDPB 2021: European Data Protection Board "EDPB Document on Response to the Request from the European Commission for Clarifications on the Consistent Application of the GDPR, Focusing on Health Research" (2021).

EDPS 2020: European Data Protection Supervisor, "A Preliminary Opinion on Data Protection and Scientific Research" (2020).

WP29 2010: Article 29 Working Party, "Opinion 1/2010 on the Concepts of 'Controller' and 'Processor'"(WP 169, 2010).

WP29 2013: Article 29 Working Party, “Opinion 03/2013 on Purpose Limitation” (WP 203, 2013).

WP29 2014: Article 29 Working Party, “Opinion 06/2014 on the Notion of Legitimate Interests of the Data Controller under Article 7 of Directive 95/46/EC” (WP 217, 2014).

WP29 2015: Article 29 Working party, “Annex – Health Data in Apps and Devices” (2015).

WP29 2015: Article 29 Working Party, “Opinion 1/2015 on Privacy and Data Protection Issues Relating to the Utilisation of Drones” (WP 231, 2015).

WP29 2018: Article 29 Working Party, “Guidelines on Consent under Regulation 2016/679” (WP 259, 2018).

WP29 2018: Article 29 Working Party, “Guidelines on Transparency under Regulation 2016/679” (WP 260, 2018).

Swedish Preparatory Work

Prop. 2002/03:50 Etikprövning av forskning.

Prop. 2007/08:44 Vissa etikprövningsfrågor m.m.

Prop. 2017/18:105 Ny dataskyddslag.

Prop. 2017/18:298 Behandling av personuppgifter för forskningsändamål.

Prop. 2018/19:165 Etikprövning av forskare – tydligare regler och skärpta straff.

SOU 2017:39 Ny dataskyddslag. Kompletterande bestämmelser till EU:s dataskyddsförordning.

SOU 2017:50 Personuppgiftsbehandling för forskningsändamål.

SOU 2018:36 Rätt att forska – Långsiktig reglering av forskningsdatabaser.

Reports

Information Commissioner’s Office (ICO), Report RFA0627721 Provision of Patient Data to DeepMind (2017).

OECD, “Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development” (2015).

Swedish Ethical Review Authority, ”Vägledning om etikprövning av forskning på människor” (2023).

Swedish Research Council, “Good Research Practice” (2017).

News Articles & Web Pages

Dermalyser. “About” and “Dermalyser,” from https://www.aimedtech.com.

Deversify. “About us” and “R&D,” from https://deversify.com.

Herper, M. “Surprise! With $60 Million Genentech Deal, 23andMe Has a Business Plan,” Forbes (Jan 6, 2015). https://www.forbes.com/sites/matthewherper/2015/01/06/surprise-with-60-million-genentech-deal-23andme-has-a-business-plan/?sh=148321422be9

Table of Cases

Case C-101/01, Criminal proceedings against Bodil Lindqvist.
ECLI:EU:C:2003:596.

Case C-212/13, František Ryneš v Úřad pro ochranu osobních údajů.
ECLI:EU:C:2014:2428.

Case C-322/88 Salvatore Grimaldi v Fonds des maladies professionnelles.
ECLI:EU:C:1989:646.

Case C-345/17, Sergejs Buivids v. Datu valsts inspekcija.
ECLI:EU:C:2019:122.

Case C-434/16, Peter Nowak v Data Protection Commissioner.
ECLI:EU:C:2017:994.

Case C-582/14, Patrick Breyer v. Bundesrepublik Deutschland.
ECLI:EU:C:2016:779.

ECtHR Case of Vukota-Bojic v. Switzerland (application number 61838/10).