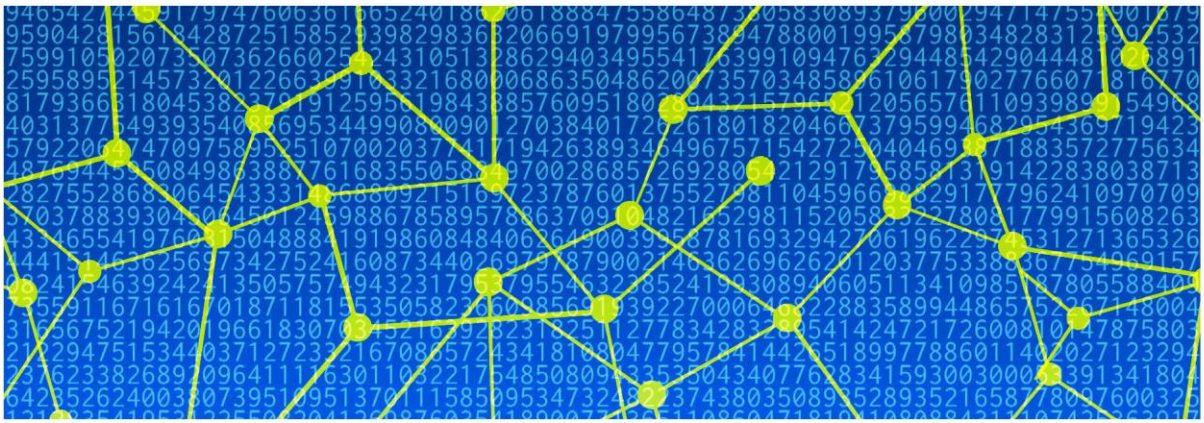

EUROPEAN DATA PROTECTION SUPERVISOR

A Preliminary Opinion on data protection and scientific research

EDPS

6 January 2020

The European Data Protection Supervisor (EDPS) is an independent EU authority, responsible under Article 52(2) of Regulation 2018/1725 'With respect to the processing of personal data… for ensuring that the fundamental rights and freedoms of natural persons, and in particular their right to data protection, are respected by Union institutions and bodies', and under Article 52(3)'…for advising Union institutions and bodies and data subjects on all matters concerning the processing of personal data'. Under ***article 58(3)(c)*** *of Regulation 2018/1725, the EDPS shall have the power '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'.*

Wojciech Wiewiorówski was appointed Supervisor on 5 December 2019 for a term of five years.

Executive Summary

Scientific research depends on the exchange of ideas, knowledge and information. Where it involves the processing of data concerning people in the EU, scientific research is subject to the applicable rules including the General Data Protection Regulation and Regulation 1725/2018 for EU institutions. The rules contain a special regime affording a degree of flexibility for genuine research projects that operate within an ethical framework and aim to grow society's collective knowledge and wellbeing. How this special regime should operate in practice is under discussion. Some argue that the GDPR offers too much flexibility, others that the rules threaten vital research activity.

Digitisation has made the generation and dissemination of personal data easier and cheaper than ever and transformed how research is carried out. The boundary between private sector research and traditional academic research is blurrier than ever, and it is ever harder to distinguish research with generalisable benefits for society from that which primarily serves private interests. Corporate secrecy, particularly in the tech sector, which controls the most valuable data for understanding the impact of digitisation and specific phenomena like the dissimilation of misinformation, is a major barrier to social science research.

In the particular field of health science, medical research and clinical trials generally take place within an established framework of professional ethical standards. The interaction between this framework and the GDPR is being discussed within the European Data Protection Board.

The special regime applies the usual principles such as lawfulness, purpose limitation and data subject rights, but permits some derogations from controller obligations. This includes the presumption of compatibility of processing for scientific research purposes of data collected in commercial and other contexts, provided appropriate safeguards are in place. This flexibility is afforded on the assumption that research occurring within a framework of ethical oversight serves, in principle, the public interest. The accountability principle therefore key, as it requires controllers to assess honestly and manage responsibly the risks inherent in their research projects. Such risks can be very high where, for example, processing sensitive data on health or political or religious views. Consent as a legal basis for processing must be freely-given, specific, informed and unambiguous. This differs conceptually and operationally from 'informed consent' of human participants in research. Such 'informed consent' may still serve as a safeguard in cases where consent is not appropriate as a data processing legal basis.

Scientific research serves a valuable function in a democratic society to hold powerful players to account, and this has grown in importance with the concentration of control over information flows in the hands of a few private global companies. Data protection obligations should not be misappropriated as a means for powerful players to escape transparency and accountability. Researchers operating within ethical governance frameworks should therefore be able to access necessary API and other data, with a valid legal basis and subject to the principle of proportionality and appropriate safeguards.

We recommend intensifying dialogue between data protection authorities and ethical review boards for a common understanding of which activities qualify as genuine research, EU codes of conduct for scientific research, closer alignment between EU research framework programmes and data protection standards, and the beginning of a debate on the circumstances in which access by researchers to data held by private companies can be based on public interest.

TABLE OF CONTENTS

1. Introduction 5

2. Research today 6

2.1 DIGITISATION 6
2.2 ACADEMIA AND THE COMMERCIAL SECTOR 7
2.3 BEHAVIOURAL EXPERIMENTS 8
2.4 CORPORATE SECRECY AS A BARRIER TO RESEARCH 9

3. The notion of scientific research 9

3.1 RESEARCH 9
3.2 SCIENTIFIC RESEARCH 10
3.3 DISTINGUISHING RESEARCH FROM ACADEMIC EXPRESSION 10
3.4 COPYRIGHTED TEXT AND DATA 11
3.5 SCOPE OF THE SPECIAL DATA PROTECTION REGIME FOR SCIENTIFIC RESEARCH 11

4. Governance and policy of research in the EU 12

4.1 PRINCIPLES OF THE EU'S RESEARCH POLICY AND FACILITATION OF DATA SHARING . 12
Open AIRE 12
European Open Science Cloud 12
Public Sector Information 12
4.2 ETHICAL STANDARDS 13
History 13
Consent and oversight 13

5. Health science 14

5.1 MEDICAL RESEARCH 14
Consent 14
5.2 CLINICAL TRIALS 15
EU Clinical Trials Regulation 15
Informed consent 15
Use of clinical trials data by other research institutions 15
Clinical trials and GDPR 16

6. Scientific research and the GDPR: Selected issues 16

6.1 PRINCIPLES OF DATA PROTECTION 16
6.2 SPECIAL REGIME FOR SCIENTIFIC RESEARCH 16
Scalability of obligations, other norms and coupling of information 16
Article 89 17
Special categories of personal data 17
Restricted flexibility of the special regime 18
6.3 CONSENT AS A LEGAL BASIS FOR DATA PROCESSING 18
6.4 RIGHT TO INFORMATION 20
6.5 DECEPTION, INFORMED CONSENT AND THE RIGHT TO INFORMATION 20
6.6 DEROGATIONS POSSIBLE UNDER THE SPECIAL REGIME FOR SCIENTIFIC RESEARCH ... 21
6.7 PURPOSE LIMITATION AND THE PRESUMPTION OF COMPATIBILITY 22
Purpose limitation principle 22
Presumption of compatibility for research purposes 22
Lawfulness and purpose limitation 22
6.8 PUBLIC INTEREST AS A BASIS FOR LAWFUL PROCESSING 23
6.9 STORAGE LIMITATION 23
6.10 ACCOUNTABILITY 24

7. Recommendations 24

7.1 DPAS AND ETHICAL REVIEW BOARDS 25
7.2 EU CODES OF CONDUCT AND CERTIFICATION FOR RESEARCH ACTIVITIES 25
7.3 EU RESEARCH FRAMEWORK PROGRAMMES AND DATA PROTECTION STANDARDS 26
7.4 DEBATE ON A PUBLIC INTEREST GROUND FOR SCIENTIFIC RESEARCH 26
8. Conclusion 26
NOTES 28

THE EUROPEAN DATA PROTECTION SUPERVISOR,

Having regard to the Treaty on the Functioning of the European Union, and in particular Article 16 thereof,

Having regard to the Charter of Fundamental Rights of the European Union, and in particular Articles 7 and 8 thereof,

Having regard to Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)[1],

Having regard to Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC[2], in particular Articles 42(1), 57(1)(g) and 58(3)(c) thereof,

Having regard to Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA[3],

HAS ADOPTED THE FOLLOWING OPINION:

1. Introduction

The advancement of generalisable knowledge is a priority for the EU[4]. Data protection rules, as a set of norms with general application, cannot be isolated from the governance structures of specific domains and disciplines. The General Data Protection Regulation (GDPR) assigns to scientific research a special regime, but there have been few guidelines or comprehensive studies on the application of data protection rules to research[5]. The previous 1995 Data Protection Directive (Directive 95/46/EC) had permitted Member States to adopt legislation specifying further the regime for data processing for research purposes, and the GDPR also allows derogations to be introduced by EU or Member State law, with the result of a patchwork of safeguards. Perceptions of the impact of the GDPR vary: some claim it is a barrier to research, others that it has little impact, still others that it represents a loophole for emptying data subjects' rights[6]. Regulation 2018/1725 governing data protection in EU institutions and bodies largely replicates the GDPR's provisions in this area, and the analysis and recommendations in this Preliminary Opinion may be generally regarded as relevant for both data processing under both regulations[7].

In the field particularly of medical research, personal data are processed 'on an unprecedented scale'[8]. The distinction between, on the one hand, genuine research for the common good and, on the other, research which serves primarily private or commercial ends, has become ever more blurred. There is a thriving market for direct-to-consumers genetic testing services offering to predict medical risk factors and to reveal ancestry or genealogy; this itself is not research, but a strategy for collecting data based on consent that can be subsequently further used for research or other purposes including law enforcement. At the same time, there is the suspicion of data protection being enlisted to escape accountability where, on the pretext of safeguarding the rights of others, inferred personal data is conflated with intellectual property,

offering a shield for corporate secrecy. There is also some (understandable) confusion regarding consent, which is a principle of both data protection and research involving human participants. There is a complex interplay between the obligations of the controller (responsible for personal data processing), of the person or entity responsible for the research (the 'sponsor' in the context of clinical trials) and the person carrying out the actual investigations, who depending on the circumstances could be a separate controller, joint controller and/or processor[9].

The EDPS considers that respect for personal data is wholly compatible with responsible research. Data protection is intended to serve as a safety net for individuals whose data are needed to support science: personal data enable better understanding of diseases, the development of new therapies and generally improvement of quality of life[10]. The concept of research, however, is very broad. This Preliminary Opinion builds on the work of the EDPB and its predecessor the Article 29 Working Party to promote a more informed discussion between the research and data protection communities.

The document is structured as follows: first, we sketch out the landscape of scientific research in today's digital age and the issues which arise (section 2). Second, we aim to narrow down what we understand by scientific research in the GDPR (section 3). Third, we outline the wider governance framework for research in the EU within which data protection is situated, particular as regards clinical trials (sections 4 and 5). Fourth, we present a preliminary analysis of some key principles of the special regime for data processing for the purposes of scientific research as set down in the GDPR (section 6). This includes in particular the notion of consent, the presumption of compatibility and derogations to data subject rights. Finally, we point to a number of areas for further consideration (section 7).

Although this Preliminary Opinion is restricted to data protection and scientific research, it should be noted that there are no agreed or precise boundaries between scientific research, other forms of research in the humanities and the arts, and marketing research. The GDPR itself applies a special data protection regime for archiving purposes in the public interest, historical research and statistics, as well as for scientific research. We wish to stress that this analysis is preliminary, not comprehensive, and we encourage constructive criticism on the content and on aspects which may be missing. Most of the follow-up to this document is likely to come from the European Data Protection Board (EDPB), particularly on the key issues and principles such as consent, retention and secondary use.

2. Research today

2.1 Digitisation

Digitisation has transformed research. The cost of data processing and storage continues to decrease, processing power increases and sensors and connected devices proliferate. Researchers, particularly in medical research, often work in large collaborative networks and need to exchange large volumes of data at great speed across borders[11]. In the online environment researchers may have limited direct contact with participants, and large-scale genomic databases are developed for use by multiple researchers over long periods.

The extraordinary reach of a small number of very powerful global technology companies has involved subsidising and sponsoring academic research on an unprecedented scale. They are able to attract talent to work in their industry, typically on condition of signing non-disclosure agreements, which in the past might have remained within academia[12]. These private companies control gigantic databases of personal information gathered through systematic and multifarious monitoring of people's activity while connected to the internet. Most data in the

world may well now be held by these companies, whereas in the past they would have been created by academics or governments[13]. These data can provide a basis for independent research: for example, a 2017 study which analysed social media data to test hypotheses about markers of depression was based on consent with approval from an ethics committee and a declaration of source of funding[14].

2.2 Academia and the commercial sector

Research is clearly no longer the preserve of academia, if indeed it ever was[15]. The interface, including the exchange of data, between research organisations and the wider research ecosystem is highly complex. Scientific publishers, designers and developers, entrepreneurs, commercial, governmental and non-profit funding sources in the commercial, governmental and non-profit sectors all have a stake. Also playing a role are big data analytics firms and cloud service providers, and platforms and apps whose business models involve the accumulation and monetisation of as much data as possible. Then there are direct-to-customer genetic testing companies, the apps and websites which offer services for interpreting the raw genetic data generated by these tests, and the pharmaceutical and med-tech companies who pay enormous amounts for access to the data. This has been described by one medical journal as a 'wild west environment' of data sharing, lacking the rigorous regulation of the health care context[16].

The intertwining of academia and the commercial sector can be seen in several ways. Firstly, funding: large companies and particularly technology companies subsidise and sponsor vast amounts of academic research, with funding often contingent on the signature of non-disclosure agreements, casting doubt on the integrity, impartiality and credibility of the research[17]. Without transparency and standards, such research is susceptible to being viewed as covert form of corporate lobbying[18].

Secondly, these companies compete fiercely to attract talent often directly from universities. Amazon was reported to have hired 150 PhD economists over five years to work on user growth, profitability and platform design[19]. These employees have a limitless store of privately-held data and the insights are covered by commercial secrecy.

Thirdly, traditional research institutions and public bodies often cooperate or enter partnership agreements with technology companies. The UK National Health Service was found to have given Google and its AI Company DeepMind access to a trove of sensitive information on 1.6 million patients including HIV status, mental health history and abortions without the patients being properly informed[20]. Despite the public reassurances that the patients' data were never 'linked to Google products or services or used for any commercial purposes in any way', there are concerns that such companies lack incentives to treat sensitive health-related data responsibly[21].

Fourth, as revealed by the UK Information Commissioner's Office investigation into Cambridge Analytica, academic studies and the commercial enterprises set up by academics can become inextricably entangled, sharing the use of university equipment for example[22]. Cambridge Analytica's targeting techniques originated in the work of academics of the Psychometrics Centre at Cambridge University[23]. There was evidence of a 'close working relationship between Facebook and individual members of the research community', while the Psychometric Centre used Facebook data for psychometric testing through the development of an online quiz and developing personality profiles[24].

Fifth, there is the increasingly lucrative market elsewhere for genetic testing services, including in Europe in spite of the varying regulatory regimes across the Member States including effective bans at least two cases[25]. Providers of these services seem to operate on the

presumption of 'data ownership' and 'broad consent' (see Section 5 below) to sell to researchers access to their data, thus becoming in effect indispensable intermediaries between researchers and their research subjects[26]. One US-based start-up, for example, offers whole genome sequencing for 'free' if they agree to allow their genetic data to be shared with medical researchers or pharmaceutical companies, and 'rewards' customers with more detailed analysis if they answer intimate questions and if third parties wish to access those data[27]. The international Personal Genome Project, launched in 1990, claims to rely on 'open access consent', where participants agree to share their genetic and genomic data in a fully open access database online[28] for scientific research purposes and in effect for any other use that might be made of the publicly available database. These techniques for attracting individuals to provide sensitive personal information are susceptible to ethical concerns with how the information will be used, as well with the context in which participants are asked to consent[29].

2.3 Behavioural experiments

Large tech companies typically have a dedicated 'research arm'. Mental health experts and psychologists have been employed to roll out 'persuasive design' to induce addiction to their devices and software. In 2018, 50 psychologists wrote protested protest against 'the unethical practice of psychologists using hidden manipulation techniques to hook children on social media and video games 'hidden manipulation techniques'[30].

The Facebook Research home page openly states that their 'Human Computer Interaction (HCI) and User Experience (UX) researchers seek to deeply understand and improve the experiences of the over 2 billion people around the world who use Facebook every month.' The company has offered e-rewards to children if they allowed its 'Research Program' to run in the background of all applications on their devices to monitor usage habits, even offering a VPN application in an apparent attempt to deceive them into believing they had secure communications channel[31].

In 2014 (as mentioned above), Facebook conducted an experiment on 'emotional contagion' via social media networks. The company's core data science team published their research stating that 'emotions expressed by others on Facebook influence our own emotions, constituting experimental evidence for massive-scale contagion via social networks'[32]. The study was claimed to be 'internal development research' and consistent with Facebook's Data Use Policy, 'to which all users agree prior to creating an account on Facebook, constituting informed consent for this research'. There are many similar examples.

- OkCupid, the online dating app, in 2014 announced that it 'took pairs of bad matches (actual 30% match) and told [users] they were exceptionally good for each other (displaying a 90% match)'[33]. In 2016, an OkCupid study publicly released a dataset of nearly 70 000 users, including usernames, age, gender, location, what kind of relationship (or sex) they were interested in, personality traits, and answers to thousands of profiling questions used by the site[34].

- Amazon admitted to be 'constantly experimenting' in response to criticism following reports of their monetising of baby registries with deceptive advertisements[35].

- An AI education company inserted 'social-psychological interventions' into one of its commercial learning software programmes affecting 9000 students without seeking their consent.

- A university created a lab which is teaching the art of persuasion in technology with the explicit aim of creating 'insight into how computing products can be designed to change what people believe and what they do'[36].

Such human behavioural experimentation is, according to the AI Now Institute, 'rampant'[37]. Our analysis in this Preliminary Opinion will attempt to explain why, in our view, these practices would not qualify as scientific research under the GDPR.

2.4 Corporate secrecy as a barrier to research

Democratic societies rely on independent researchers in the social sciences being able to investigate and explain the role of these companies in determining how information is processed and flows around Europe and the world[38]. Research includes the examination of the important phenomenon of digitisation itself. The ability of independent researchers to investigate the role of big technology companies is part of a wider necessary accountability framework alongside the independent regulatory oversight. However, the widely reported corporate secrecy which characterises the biggest technology companies is a barrier to scrutiny by such researchers, similar to reported refusals to submit to independent audits by government agencies[39]. Following the Cambridge Analytica scandal - which stemmed from activities outside the normal ethical framework - academics seem to have more difficulty interrogate information flows. For instance, Facebook in 2018 restricted access to Application programming interface data ('API' - that is, data on how they process requests for Facebook data from remote applications) and sensitive targeting data. This prevented researchers from performing network analysis or finding the connections between accounts in order to understand the dissemination of conspiracy theories, hate speech and disinformation.

Resistance to greater transparency and accountability is justified on questionable grounds of data protection. There are concerns that the references to fundamental rights in Code of Conduct on Disinformation could be a cover for similar attempts to avoid scrutiny[40]. Independent researchers called on the major platforms to allow access in analysable format to advertising archive API in order to monitor election influence and the spread of disinformation and to hold powerful players to account[41]. In December 2019, the European Advisory Committee of Social Science One, an initiative designed to provide academics with access to digital platforms data, reported that Facebook had not provided any adequate access with the result that researchers are 'left in the dark' and unable to 'make full informed contributions to discussions' over 'the role and responsibilities of platforms[42].'

All data processing, including the sharing of data with third parties, involves a degree of risk that needs to be carefully assessed and managed. Some risks may be so high that the processing should not take place at all. By contrast, privacy policies for web-based services, including for the major platforms, are notoriously vague: companies appear to give themselves maximum latitude in determining how, and for what purpose, they can use the personal data collected, and with whom they can share it[43]. The same policies often also state the possibility to share information for research purposes. There remain few examples, however, of private companies making the data they hold available. It would appear therefore that the reluctance to give access to genuine researchers is motivated no so much by data protection concerns as by the absence of business incentive to invest effort in disclosing or being transparent about the volume and nature of data they control[44].

3. The notion of scientific research

3.1 Research

There is no universally agreed definition of research or scientific research. Reputable definitions of research tend to emphasis systematic activity, including the gathering and analysis of data, which increases the stock of understanding and knowledge and their application[45]. The European Commission has defined the objectives of the EU's research and

innovation policies to be 'opening up the innovation process to people with experience in fields other than academia and science', 'spreading knowledge as soon as it is available using digital and collaborative technology' and 'promoting international cooperation in the research community'[46].

3.2 Scientific research

Scientific research applies the 'scientific method' of observing phenomena, formulating and testing a hypothesis for those phenomena, and concluding as to the validity of the hypothesis[47]. Within individual disciplines, researchers may profess to serve a general aim, such as the personalisation of individual treatment through closer understanding of the individual and his/her susceptibility to a given condition[48]. The conduct of research must allow testing of hypotheses, with both the conclusion and the reasoning transparent and open to criticism. Openness and transparency help distinguish between science and pseudo-science.

The EU's 2019 Copyright Directive (Directive (EU) 2019/790) considers scientific research to cover 'both the natural sciences and the human sciences', and distinguishes between not-for-profit and public interest bodies and organisations operating under commercial influences:[49]

> *Due to the diversity of such entities, it is important to have a common understanding of research organisations. They should for example cover, in addition to universities or other higher education institutions and their libraries, also entities such as research institutes and hospitals that carry out research. Despite different legal forms and structures, research organisations in the Member States generally have in common that they act either on a not-for-profit basis or in the context of a public-interest mission recognised by the State. Such a public-interest mission could, for example, be reflected through public funding or through provisions in national laws or public contracts. Conversely, organisations upon which commercial undertakings have a decisive influence allowing such undertakings to exercise control because of structural situations, such as through their quality of shareholder or member, which could result in preferential access to the results of the research, should not be considered research organisations for the purposes of this Directive.*

3.3 Distinguishing research from academic expression

The European Court of Human Rights has referred to 'academics' freedom to express freely their opinion about the institution or system in which they work and freedom to distribute knowledge and truth without restriction'[50]. The Court cited Recommendation 1762 (2006) of the Parliamentary Assembly of the Council of Europe concerning the protection of academic freedom of expression. According to this Recommendation:

> *Academic freedom in research and in training should guarantee freedom of expression and of action, freedom to disseminate information and freedom to conduct research and distribute knowledge and truth without restriction.*

The GDPR sets down a regime set out for data processing for 'journalistic purposes and the purposes of academic, artistic or literary expression' (Article 85)[51]. The scope of exemptions from GDPR provisions is here broader than the special regime for scientific research[52]. We would argue that the processing of personal data for the purposes of 'academic expression' implies: (1) processing directly linked to the freedom of academics to disseminate information, (2) their freedom to distribute knowledge and truth without restriction, such as with publications, dissemination of research results, and (3) the sharing of data and methodologies with peers and exchanges of views and opinions.

There has been some debate as to the distinction between biomedical research, academic research in the humanities and social sciences, and the Article 85 provisions. The distinction may not be always easy to apply in practice. Some overlaps may occur where scientific research conducted by universities and academic institutions could fall, to some limited extent, within the scope of both regimes. That said, the broader exemptions for academic expression cannot be interpreted as a means of justifying the circumvention of the safeguards that Article 89 requires for scientific research. As with all other derogations under the GDPR, further tailored derogations to these scientific research activities should only happen where strictly necessary.

3.4 Copyrighted text and data

As mentioned above, the new EU Copyright Directive provides for an exception to the copyright rules allowing researchers linked to a university or another research organisation to analyse text and data on a large scale, as long as the access is lawful and the research is carried out with a public interest goal[53]. This type of analysis is referred to as text and data mining: any automated analytical technique aiming to analyse text and data in digital form in order to generate information such as patterns, trends and correlations. It is therefore relevant for big datasets and for the training of artificial intelligence systems[54].

The exemption may in principle encompass research organisations in a public-private partnership, but excludes individual researchers and organisations controlled by a private undertaking. It therefore affirms the principle that genuine research using massive data sets accumulated in the digital economy must be carried out with a public interest aim and within institutional structures for accountability[55].

3.5 Scope of the special data protection regime for scientific research

The special regime in the GDPR for scientific research is composed of specific derogations from certain controller obligations plus a specific provision (Article 89) requiring appropriate safeguards. It thus reflects a clear intention to adapt data protection rules to the specific circumstances and public interests served by research activities. The fundamental rights at stake along with the right to privacy and the right to the protection of personal data, including the right to the integrity of the person and the freedom of the arts and sciences, should not be viewed as in conflict. Rather, the objective should be to seek a 'fair balance' between individual rights and other interests.

Under the GDPR, the role of research is understood to provide knowledge that can in turn 'improve the quality of life for a number of people and improve the efficiency of social services'[56]. The GDPR assumes a broad conception of research, including technological development, fundamental and applied research and privately funded research and 'studies conducted in the public interest in the area of public health'[57]. It also recommends that data processing 'take into account the EU's objective under Article 179(1) TFEU of achieving a European Research Area'[58]. Therefore, not only academic researchers but also not-for-profit organisations, governmental institutions or profit-seeking commercial companies can carry out scientific research.

It is a common assumption that scientific research is beneficial to the whole of society and that scientific knowledge is a public good to be encouraged and supported. This translates into a form of 'social contract', rooted as such in trust. In this context, where trust plays such a crucial role, performing an activity deemed to be research cannot be a *carte blanche* to take irresponsible risks. From a data protection viewpoint, the principles of necessity and proportionality are essential. For a controller to simply claim to process data for the purposes of scientific research is not sufficient[59]. The Article 29 Working Party, in its guidelines on consent, understood scientific research as a 'research project set up in accordance with relevant

sector-related methodological and ethical standards'[60]. Under this approach, only scientific research performed within an established ethical framework would therefore qualify as activities falling within the special data protection regime[61].

For the purposes of this Preliminary Opinion, therefore, the special data protection regime for scientific research is understood to apply where each of the three criteria are met:

1) personal data are processed;

2) relevant sectoral standards of methodology and ethics apply, including the notion of informed consent, accountability and oversight;

3) the research is carried out with the aim of growing society's collective knowledge and wellbeing, as opposed to serving primarily one or several private interests[62].

4. Governance and policy of research in the EU

4.1 Principles of the EU's research policy and facilitation of data sharing

Article 13 of the Charter of Fundamental Rights requires that 'The arts and scientific research shall be free of constraint. Academic freedom shall be respected.' Article 179(1) TFEU states:

> *The Union shall have the objective of strengthening its scientific and technological bases by achieving a European research area in which researchers, scientific knowledge and technology circulate freely, and encouraging it to become more competitive, including in its industry, while promoting all the research activities deemed necessary by virtue of other Chapters of the Treaties.*

The EU therefore encourages sharing and re-use of research data. Public funding is directed at research projects on condition that the results are disseminated and shared with the public[63].

Open AIRE

In 2009, the Open Access Infrastructure for Research in Europe (OpenAIRE) project was launched, aiming to build a participatory infrastructure for the Commission's Pilot for Open Access to Research Information so as to develop open interfaces to exchange research information between research data repositories[64].

European Open Science Cloud

By 2020, the European Open Science Cloud is planned as part of the European Cloud initiative[65]. This will be made possible by developing high performance computing systems which can store and process the large amount of scientific data from EU projects. It will function as a free and open virtual environment to store, manage, analyse and reuse research data across borders and scientific disciplines[66]. It is currently in the testing phase in which several field-specific projects are carried out to explore the possibilities and challenges of such a large-scale infrastructure[67].

Public Sector Information

The Public Sector Information Directive requires public sector bodies to allow reuse of all public information[68]. The latest revision of the Directive, adopted in July 2019, requires Member States to 'support the availability of research data' with measures to make 'publicly funded research data openly available'[69]. There are limits to this openness, with the revised Commission Recommendation on access to and preservation of scientific information referring to the principle that research data should be 'as open as possible, as closed as necessary'[70].

This scientific information includes not only peer-reviewed publications but also the data on which the conclusions are based so that other researchers can replicate the method followed or to build on it[71]. It allows the combination of data across multiple studies and builds on rather than replicates existing research, meaning that data may be used for reasons very different from those of the original collection.

4.2 Ethical standards

History

Ethical standards for research have evolved governing primarily medical experiments on humans. They have been adapted generally to any research using human subjects. The Nuremberg Code was probably the first example of an ethical code in modern times, formulated in reaction to the medical experimentations conducted in Nazi concentration camps[72]. Later, in 1964, the World Medical Association's Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects (last amended in 2013) included within its scope 'research on identifiable human material and data', and prescribed that 'Every precaution must be taken to protect the privacy of research subjects and the confidentiality of their personal information.' It further stipulated, 'For medical research using identifiable human material or data, such as research on material or data contained in biobanks or similar repositories, physicians must seek informed consent for its collection, storage and/or reuse. There may be exceptional situations where consent would be impossible or impracticable to obtain for such research. In such situations the research may be done only after consideration and approval of a research ethics committee'[73].

The United States, following the Declaration of Helsinki, has since 1981 applied 'the Common Rule' as an ethical standard governing biomedical and behavioural research involving human subjects. The Common Rule governs oversight of federally funded human research and is incorporated in the 1991 revision to the U.S. Department of Health and Human Services Title 45 CFR 46 (Public Welfare). It includes requirements for researchers' obtaining and documenting informed consent and Institutional Review Boards with additional safeguards for certain vulnerable research subjects namely pregnant women, foetuses, prisoners and children. Many academic journals require Common Rule compliance for all research, including where privately funded, and there are several state laws on human subject research[74].

The Council of Europe in 1997 adopted the Convention on Human Rights and Biomedicine ("the Oviedo Convention") which stressed that the human being has primacy over the sole interest of society or science[75]. It states that research on a person can be performed only subject to conditions such as the lack of an alternative of comparable effectiveness, the approval of the research project by the competent body after independent examination of its scientific merit, and ensuring the person undergoing research has been informed of the safeguards and her rights[76].

Consent and oversight

There are normally two basic components to these ethical standards:

1. informed consent; and
2. independent ethical oversight.

These represent the essential safeguards against researchers determining by themselves and in isolation the permissibility of an experiment.

The first component requires researchers to obtain the informed consent of all human participants in any of the research project. This is included in the Declaration of Helsinki[77] and is intended to enact respect for research participants. Researchers are expected to disclose information about a study's purpose, risks, procedures as well as measures in the case of harms resulting from participation.

The second component requires research involving human participants to be reviewed by independent ethics committees or Institutional Review Boards who consider whether the research is ethical, lawful and provides appropriate safeguards. Universities and other institutions have accordingly set up research committees and codes of practice[78]. Historically, national research ethics committees have examined the rights of the research subjects and wider societal implications in thorough reviews of individual research project applications[79]. They can also contribute to convergences of the national approaches through collaboration and exchange of good practices, for example via the European Network of Research Ethics Committees[80].

These standards are even more essential now that vast quantities of data are available. There is growing awareness that what has become possible through digital technology is not necessarily sustainable or justifiable: for example, consider controversy over studies claiming analysis of faces could reveal sexual orientation or a tendency toward criminality, and over the claim by the researcher in China to have applied CRISPR-Cas9 DNA modification technology on the foetuses of twin girls born in November 2018[81].

5. Health science

5.1 Medical research

Human dignity and the right to the integrity of the person are recognised in Articles 1 and 3 of the Charter of Fundamental Rights of the European Union. Medical research on humans (also known as biomedical research or experimental medicine, including 'bench science' and applied research), is strictly subject to ethical standards and controls. Under Article 3(2)(a) of the Charter, the 'free and informed consent of the person concerned' must be respected in the field of biology and medicine.

In the health sector there is, it is argued, an 'ethical and scientific imperative' to share personal data for research purposes[82]. The EU like governments elsewhere in the world promote the public sharing of anonymised clinical trial documents[83], though techniques are not standardised.

Consent

EU Member States generally require prior informed consent from the participant in a research project for the processing of health data, but in emergency situations national provisions vary across the EU[84]. Researchers, particularly in biobanking, increasingly rely on 'broad consent' to the use of data for further scientific research projects that are unknown at the time of collection, on the grounds that the risks are very low[85]. For personal genome testing, 'tiered consent', where participants are invited to select from a set of options, has been proposed[86]. 'Dynamic consent', where participants are asked to consent to different activities over time via an IT interface, has been trialled in the field of biobanks[87]. In critical care situations, such as where a person is unconscious and his/her relatives are not contactable, consent to observational research – where data is derived from the patient record or tissue samples while he/she is receiving care - cannot be obtained from either the patient or his/her proxy[88]. In such cases ethical questions regarding, for instance, the appropriateness of deferred consent or

consent from an independent physician, are the subject of discussion within the medical research community.

5.2 Clinical trials

EU Clinical Trials Regulation

Clinical trials contribute to wider medical research, and the only EU legal instrument governing any specific area of scientific research is the Clinical Trials Regulation[89], although there are also other relevant general instruments governing public sector information, access to public documents, copyrighted text and data and, of course, the GDPR. The Regulation aims to harmonise the applicable rules by introducing an authorisation procedure based on a single submission via a single EU portal, an assessment procedure leading to a single decision, rules on the protection of individuals, and informed consent and transparency requirements. The Regulation entered into force in June 2014 but before it can be applicable there must be a fully functional EU clinical trials portal and database to be subject to an independent audit and a confirmation notice published by the European Commission – currently estimated for 2020.

The Regulation defines clinical trials as a subset of clinical studies, meaning 'any investigation in relation to humans … with the objective of ascertaining the safety and/or efficacy of [...] medicinal products.' The Regulation expands on established ethical norms by setting down obligations for those responsible for the clinical trial, referred to as sponsor(s) (Recital 59), which may be an individual, company, institution or organisation (Article 2(2)(14)).

Informed consent

The Regulation details requirements for consent, which is defined as[90]:

> *a subject's free and voluntary expression of his or her willingness to participate in a particular clinical trial, after having been informed of all aspects of the clinical trial that are relevant to the subject's decision to participate or, in case of minors and of incapacitated subjects, an authorisation or agreement from their legally designated representative to include them in the clinical trial.*

Such informed consent must be[91]:

> *written, dated and signed by the person performing the interview ... and by the subject or, where the subject is not able to give informed consent, his or her legally designated representative after having been duly informed ...*

The research participant must receive understandable information during a prior interview and should be able to ask questions at any time. In addition, the participant must be given adequate time to consider his or her decision[92]. All circumstances which might influence the decision to take part in a clinical trial should be taken into account, including the economic and social status of the participant or whether the person is in a situation of institutional or hierarchical dependency[93].

Use of clinical trials data by other research institutions

Universities and other research institutions may, in accordance with data protection rules, collect data from clinical trials for use – outside the protocol of the clinical trial – such as for the purposes of medical, natural or social sciences research in accordance with applicable data protection law[94] (Article 28(2) of the Regulation). However, in such cases the research participant must give consent to use his or her data outside the clinical trial and this consent

can be withdrawn at any time[95]. A prior review on the appropriateness of such research on human data, for example on ethical aspects, is required[96].

Clinical trials and GDPR

The EDPB in January 2019 issued an opinion on the lawful grounds for data processing in the context of the Clinical Trials Regulation, which informed the revision by the European Commission in April 2019 of its Q&A on the interaction of the Regulation and the GDPR[97]. The EDPB intends to issue guidance on the 'horizontal and complex' conditions for the applicability of the 'presumption of compatibility' of further processing for archiving purposes in the public interest, scientific, historical research or statistical purposes, as provided for by the GDPR Article 5(1)(b).

6. Scientific research and the GDPR: Selected issues

6.1 Principles of data protection

The aim of EU data protection law has traditionally been to facilitate the free flow of data within the EU under common standards for lawful processing, while safeguarding the fundamental rights of individuals. The GDPR sets down six principles for collection, use, sharing and storage of personal information data processing: lawfulness, fairness and transparency; purpose limitation; data minimisation; accuracy; storage limitation; and integrity and confidentiality. To these principles is added a seventh – accountability – which means being able to demonstrate compliance with the foregoing principles.

Lawfulness, fairness, purpose limitation and the rights of access and rectification are essential elements of the right to the protection of personal data under Article 8(2) of the Charter. Any limitation of this right, according to Article 52(1) of the Charter[98]

> *must be provided for by law and respect the essence of those rights and freedoms. Subject to the principle of proportionality, limitations may be made only if they are necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others.*

6.2 Special regime for scientific research

Each of the principles under Article 5 of the GDPR apply to all data processing, including processing for research purposes. Prior to the GDPR, Directive 95/46 recognised research as an important area of public interest justifying derogations from the general rules[99]. It left data protection in the areas of health and medical research largely to Member States to legislate nationally. Research and patient groups were among the very active contributors to the negotiations of the GDPR as a single harmonised framework.

Scalability of obligations, other norms and coupling of information

Data protection obligations scale up according to the risk the processing activities pose to the individual. The data protection framework does not exist in isolation, which is why the GDPR recognises other rights and interests, including research, which may justify adjustments in, or derogations from, the general principles. In the case of scientific research, the GDPR presumes the existence of accepted and long-standing ethical and professional norms governing research on humans[100]. It also recognises the value of coupling information from registries in medical research and social sciences[101]:

> *By coupling information from registries, researchers can obtain new knowledge of great value with regard to widespread medical conditions… On the basis of registries,*

research results can be enhanced, as they draw on a larger population. Within social science, research on the basis of registries enables researchers to obtain essential knowledge about the long-term correlation of a number of social conditions.

Article 89

Article 89 of the GDPR provides for flexibility in the obligations on controllers and an emphasis on safeguards and accountability. It is worth quoting in full:

> *1. Processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner.*
>
> *2. Where personal data are processed for scientific or historical research purposes or statistical purposes, Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes.*
>
> *3. Where personal data are processed for archiving purposes in the public interest, Union or Member State law may provide for derogations from the rights referred to in Articles 15, 16, 18, 19, 20 and 21 subject to the conditions and safeguards referred to in paragraph 1 of this Article in so far as such rights are likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes.*
>
> *4. Where processing referred to in paragraphs 2 and 3 serves at the same time another purpose, the derogations shall apply only to processing for the purposes referred to in those paragraphs.*

Special categories of personal data

GDPR Article 9(2)(g) to (j) permits derogations to the prohibition of the processing of special categories of data on the basis of EU or Member State law, including for the purposes of scientific research (Article 9(2)(j) – which is a new provision in the GDPR allowing processing:

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

Member States are also able under the GDPR to enact 'further conditions, including limitations, with regard to the processing of genetic data, biometric data or data concerning health' (Article 9(4)). This is therefore a new area and requires adoption of EU or Member State law before the use of special categories of data for research purposes can become fully operational.

Restricted flexibility of the special regime

All the provisions above outline a special regime for scientific research and demonstrate that research occupies a privileged position within the GDPR. This flexibility afforded to Member States through the provisions cited above, absent harmonised EU law except in a few areas (such as for clinical trials, see Section 5.2 above), means that the full extent of this special regime is not precisely delineated[102]. Nevertheless, the special regime cannot be applied in such a way that the essence of the right to data protection is emptied out, and this includes data subject rights, appropriate organisational and technical measures against accidental or unlawful destruction, loss or alteration, and the supervision of an independent authority[103]. Personal data which are 'publicly available' - such as those collected from social media sites - are still personal data. Any limitations to fundamental rights in law are to be interpreted restrictively and cannot be abused[104]. It might be considered abusive for instance for a research organisation to interpret these special provisions in the GDPR as allowing the retention of personal data for indefinite periods and to deny data subjects rights to information. Further work is taking place on these questions within the EDPB and at national level.

6.3 Consent as a legal basis for data processing

Definition

The GDPR defines consent as[105]:

> *any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data concerning him or her...*

In the view of the EDPB, valid consent depends on these cumulative criteria[106]. In the case of sensitive data, consent must also be explicit (Article 9(2)(a)). The notion is the ongoing subject of examination in CJEU in several recent and pending judgments[107]. The definition is thus distinct from the definition of informed consent in clinical trials.

Freely-given

The requirement of consent to be 'freely given' implies genuine choice and control for data subjects[108]. The GDPR rules out consent as a valid legal ground for the processing of personal data in cases where there is a clear imbalance between the data subject and the controller[109]. Therefore, the use of enticements, inducements or rewards to elicit consent may call into question the extent to which such consent is 'freely-given'. Regarding these techniques in the specific context of clinical trials, the EDPB stated that the validity of consent as a legal basis could be in doubt where a participant is in a poor condition of health or belongs to a socio-economically disadvantaged group. This has led the EDPB to distinguish between the requirement for 'informed consent' under the Clinical Trials Regulation and the notion of explicit consent as a ground for processing special categories of data under the GDPR.[110]

Specific, informed and unambiguous

The requirement that consent be 'specific' aims to ensure a degree of user control and transparency for the data subject[111]. The consent of the data subject should be given in relation to one or more specific purposes and the data subject should have a real choice in relation to each of them[112]. According to recent case law of the CJEU, valid consent cannot be in the form of pre-ticked boxes but must be 'an active behaviour with a clear view on the part of the data subject with a view to giving his or her consent[113].'

It has been noted above that researchers have sought to rely on 'broad consent' to the use of data. Recital 33 of the GDPR indeed acknowledges that

> *it is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose.*

Specific consent normally required under the GDPR may therefore become less appropriate in the case of collected and inferred data and especially in the case of special categories of data on which much scientific research relies[114]. Recital 33 does not however take precedence over the conditions for consent set out in Articles 4(11), 6(1)(a), 7 and 9(2)(a) of the GDPR, and it requires the controller to carefully evaluate the rights of the data subject, the sensitivity of the data, the nature and purpose of the research and the relevant ethical standards. Therefore, when research purposes cannot be fully specified, a controller would be expected to do more to ensure the essence of the data subject rights to valid consent are served, including through as much transparency as possible and other safeguards[115].

Explicit consent and special categories of data

Processing of special categories of data is prohibited unless the data subject has given his or her explicit consent. Explicit consent is also a potential legal basis, in the context of scientific research, for automated decision making (Article 22(2)(c)) and for the transfer of personal data to a third country in the absence of an adequacy decision (Article 49(1)(a). Explicit consent, described by the EDPB as 'an express statement of consent' which can be demonstrated in the event of doubt, is required thus in situations where there may be particular risk to the rights of the data subject.

Special categories of data may be processed if the data subject has manifestly made them public. EU data protection authorities have argued that this provision has to be 'interpreted to imply that the data subject was aware that the respective data will be publicly available which means to everyone' including, in this case, researchers, and that, 'In case of doubt, a narrow interpretation should be applied, as the assumption is that the data subject has voluntarily given up the special protection for sensitive data by making them available to the public including authorities'.[116] Publishing personal data in a biography or an article in the press is not the same as posting a message on a social media page.

Able to withdraw

If consent is the lawful ground for processing, the data subject must be able to withdraw that consent at any time[117]; there is no exception to this requirement for scientific research. As a general rule, if consent is withdrawn, the controller is required to stop the processing actions concerned and, unless there is another lawful basis for the retention of those data for further processing, the data should be deleted by the controller[118].

Discussion: Consent of human participants in research and consent of data subjects

There is clear overlap between *informed consent* of human participants in research projects involving humans and *consent* under data protection law. But to view them as a single and indivisible requirement would be simplistic and misleading. Consent serves not only as a possible legal basis for the activity, it is also a safeguard - a means for giving individuals more control and choice and thereby for upholding society's trust in science.

There may be circumstances in which consent is not the most suitable legal basis for data processing, and other lawful grounds under both Articles 6 and 9 GDPR should be considered. However, even where consent is not appropriate as a legal basis under GDPR, *informed consent* as a human research participant could still serve as an 'appropriate safeguard' of the rights of the data subject. Under what conditions such *informed consent* might be deemed an appropriate safeguard is still unclear. Certainly, innovative forms of consent in research activities, like tiered and dynamic consent (see Section 5 above), are promising practices that should be further encouraged and developed.

The notion of consent in the two areas requires further discussion between the research community and data protection experts as part of a wider reflection on the role of consent and respect for individuals in the area of scientific research in the digital age.

6.4 Right to information

The principles of fairness and transparency (Article 5(1) GDPR) echo to a large extent the foundational principle of informed consent in research ethics, according to which participants should understand that they are taking part in research and what the research requires of them, without having been coerced or deceived**.** Such information may include the purpose of the research, the methods being used, the possible outcomes of the research, as well as associated demands, discomforts, inconveniences and risks that the participants may face.

Under Article 13 GDPR, where data are collected from an individual, he or she must be informed as to who is collecting, how to contact the controller and the data protection officer, for which purpose and on which legal grounds the data is processed, who will also receive the data, for how long it will be kept and how this period is determined, and whether automated decision-making is involved. This also includes receiving information on the rights available to them as well as the right to lodge a complaint with a supervisory authority[119].

Where data have not been obtained from the individual, patients and people participating in research have, in principle, the right to be fully informed that a data concerning him or her is being processed. Article 14 GDPR sets out the data subject's right to information – what information should be provided and how. The information provided should also include the categories of personal data which are processed, the source from which the data comes and whether it came from publicly accessible sources[120]. However, under Article 14(5)(b), the obligation to provide information does not apply if it 'proves impossible or would involve a disproportionate effort, in particular for processing for scientific research purposes when the conditions of Article 89 are satisfied or when this is likely to render impossible or seriously impair the achievement of the objective of that processing'.

In determining what constitutes either impossibility or disproportionate effort, Recital 62 refers to the number of data subjects, the age of the data and appropriate safeguards in place as possible indicative factors. The Article 29 Working Party further emphasised that the controller should carry out a balancing exercise to assess the effort involved to provide the information to data subjects against the impact and effects on the data subject if they are not provided with the information[121].

In case of further use for a different purpose, the participants must be informed before further processing takes place, even if the purpose is compatible[122].

6.5 Deception, informed consent and the right to information

As discussed above (section 6.3), data protection experts and the research community could further reflect on research activities where consent is not an appropriate legal basis for data

processing but 'informed consent' may be an appropriate safeguard. This is especially true where research studies involve the deception of subjects. Deception may include withholding information in the instructions to research participants, providing only limited information as to the purpose of the research or even misleading participants by providing a 'cover story' for the study to mask the actual topic of the study. In some psychology experiments known as covered research, subjects are misled about what is being tested, and this is cited as a key success factor because awareness of the exact nature of the research would alter people's behaviour[123]. In deception, subjects normally know they are observed, but do not know the real objective. Such cases of manipulation are often discouraged by ethics boards, but nevertheless is still used in some selected projects. In such cases, debriefing of the research participants and retrospective informed consent along with specific ethics approval before the start of the research are among the measures to ensure ethics compliance.

These practices appear to conflict with the right to information under data protection law. Indeed, there are no derogations to the principle of transparency under Article 13, where information from participants is collected directly by researchers. The possible derogation for the purposes of scientific research under Article 14 is not relevant for these specific cases of deception, since the derogation only applies in case of indirect collection. Clearly further analysis and discussion is needed on this question.

6.6 Derogations possible under the special regime for scientific research

Article 89(2) of the GDPR outlines the more specific conditions under which EU or Member State law may derogate from the data subject's right of access (Article 15), right to rectification (Article 16), right to restriction (Article 18) and right to object (Article 21).

The rights of access and rectification are set out in Article 8(2) of the Charter itself, and are generally considered essential components of the right to the protection of personal data. The right of access is of particular importance as it enables the data subjects to exercise the other rights provided for by data protection legislation[124]. Therefore, any derogation from these essential data subject rights must be subject to a particularly high level of scrutiny in line with the standards required by Article 52(1) of the Charter. Derogations under GDPR Article 89(2) are only possible if the conditions and safeguards required under Article 89(1) are satisfied[125].

Furthermore, under Article 89(2), derogations can be applied only 'in so far as' the rights to be derogated from are 'likely to render impossible or seriously impair the achievement of the specific purposes, and such derogations are necessary for the fulfilment of those purposes'. The EDPS has previously argued that this sets a high bar, in line with Article 52(1) of the Charter[126]. Enabling individuals to exercise their rights of access, rectification, restriction and objection undeniably requires a number of technical and organisational measures to be put in place by the controller.

Some of these technical and organisational measures may involve significant investment of human and financial resources in order to provide access and other rights to individuals. This is, however, not unique to companies or organisations involved in statistical or scientific research activities. Having to invest resources in itself does not justify derogating from the rights of individuals under Article 89(2) of the GDPR.

As far as the rights to restriction and objection are concerned, we recognise that in specific circumstances, a large number of individuals objecting to all or part of a scientific research, may have a negative effect on the representativeness and reliability of the research data, and thus on the integrity of research. By way of illustration, the withdrawal of consent or objection of individuals to certain research activities related to rare diseases may have a significant effect on, and possibly 'seriously impair', the outcome of long-term research studies. This is by no

means true in all cases, however. The scope of the derogations to the rights to restriction and objection in the field of scientific research should, therefore, remain limited to cases where the integrity of research would be compromised by the exercise of data subjects' rights.

6.7 Purpose limitation and the presumption of compatibility

Purpose limitation principle

Under the principle of purpose limitation, personal data must always be collected for specified, explicit and legitimate purposes and further processing of the same data is not permitted for purposes incompatible with the original purpose for processing[127]. By contrast, when personal data are further used for compatible purposes, 'no legal basis separate from that which allowed the collection of the personal data is required'. This is premised on 'the reasonable expectations of data subjects based on their relationship with the controller as to [the data's] further use' (Recital 50 GDPR).

Article 6(4) GDPR establishes criteria for determining the compatibility of further or secondary use of personal data, largely following the guidelines of the Article 29 Working Party[128]. Controllers are required to consider 'any link between the purposes for which the personal data have been collected and the purposes of the intended further processing' (Article 6(4)(a)) or the context in which personal data have been collected' (Article 6(4)(b)). If the controller shares or further processes the data for purposes incompatible with the original purposes then a new valid legal basis may be needed[129].

Presumption of compatibility for research purposes

Reflecting the strategic importance of reuse of data under the EU's research policies, EU data protection law, since the 1995 directive and continuing with the GDPR, has included the so-called presumption of compatibility (GDPR Article 5(1)(b)[130] according to which:

> *further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes*.

This presumption depends on the requirement in Article 89(1) to ensure appropriate technical and organisational safeguards, such as pseudonymisation and access limitations. The Article 29 Working Party furthermore argued for ensuring that the data would not be used to support measures or decisions regarding any particular individuals[131]. The presumption is not a general authorisation to further process data in all cases for historical, statistical or scientific purposes. Each case must be considered on its own merits and circumstances. But 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.

Lawfulness and purpose limitation

The notion of compatibility and the principle of lawfulness requires careful analysis. Under Article 8(2) of the Charter, the requirement of purpose specification is separate to the requirement of the data subject's consent or other legitimate basis[132]. Accordingly, in interpreting Directive 95/46, the Article 29 Working Party's guidelines on purpose limitation considered purpose specification and lawfulness to be two separate and cumulative requirements[133], meaning that any re-use of data for scientific research purposes, even with the presumption of compatibility, would still require a specific lawful ground.

Recital 50 GDPR states however that when personal data are used for secondary compatible purposes,

> *no legal basis separate from that which allowed the collection of the personal data is required*[and] … *(f)urther processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes should be considered to be compatible lawful processing operations*.

The recital thus appears to assimilate purpose specification and lawfulness in the case of reuse for the purposes of scientific research. As the recital is not accompanied by a specific provision in the main body of the GDPR, this appears not so much a blanket exemption to the separate steps set out in the Charter Article 8(2) - applicable to all circumstances - but rather advisory (hence 'should be considered to be compatible'). We would therefore argue that, in order to ensure respect for the rights of the data subject, the compatibility test under Article 6(4) should still be considered prior to the reuse of data for the purposes of scientific research, particularly where the data was originally collected for very different purposes or outside the area of scientific research. Indeed, according to one analysis from a medical research perspective, applying this test should be straightforward[134].

6.8 Public interest as a basis for lawful processing

The processing of personal data may be considered 'necessary for the performance of a task carried out in the public interest' (Article 6(1)(e) GDPR). In such cases, that public interest should be (Article 6(3) GDPR) laid down by EU or Member State law. A further exemption to the prohibition on processing special categories of data is processing that is (Article 9(2)(i)

> *necessary for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices, on the basis of Union or Member State law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy*

According to European case law, necessity and the public interest imply a 'pressing social need', as opposed to largely private or commercial advantages[135].

In the case of special categories of data, (Article 9(2)(g)), data processing must be 'necessary for reasons of substantial public interest'. Article 9(2)(j) (as stated above, Section 6.2) in principle provides for processing of special categories of data for scientific research but only on the basis of EU or Member State law. However, such laws have yet to be adopted. It is therefore difficult at present, if not impossible, to view a 'substantial public interest' as a basis for processing sensitive data for scientific research purposes.

Recently there have been calls for regulated access across the EU to personal data for research purposes that serve a public interest (e.g. to improve healthcare provision), noting the uncertainty around what counts as 'scientific research'[136]. This is an interesting area requiring further work.

6.9 Storage limitation

Personal data should be 'kept in a form which permits identification of data subjects for no longer than is necessary' (Article 5(1)(e) GDPR). The GDPR permits 'storage for longer periods' if the *sole* purpose is scientific research (or archiving in the public interest, historical research or statistical purposes). The intention of the lawmaker appears to have been to dissuade *unlimited* storage even in this special regime, and guards against scientific research

as a pretext for longer storage for other, private, purposes. If in doubt, the controller should consider whether a new legal basis is appropriate.

6.10 Accountability

The accountability principle in the GDPR requires controllers and processors to take responsibility for their processing activities and for having in place measures and records to demonstrate compliance. It is closely tied to the GDPR's risk-based approach, according to which data protection should be proportionate to the risks for the data subjects entailed in the data processing. The new framework sought to cut red tape, avoiding many *ex-ante* obligations, and to hold those handling personal data accountable on the basis of how they structure and ensure adequate levels of protection. The higher the risk to individuals, the higher the level of protection, therefore of positive obligations, and safeguards to implement for data controllers and processors.

Scientific research often involves the processing and sharing of sensitive types of personal data of the people involved. As such it would reasonably be considered a high-risk data processing activity according to the GDPR. There is a framework for accessing public sector information (see section 4.1 above), but not for private companies to provide researchers access to their data. Nevertheless, many companies including the big platforms state in their privacy policies that they may share data with researchers. As discussed above (section 2.4), this access can contribute to greater accountability where independent researchers are able to examine the role of powerful private players in determining how personal data is processed and information flows. There is however a reluctance to facilitate such access.

Data protection rules are a framework for, not an obstacle to, proportionate disclosure of information to researchers, where there is a valid legal basis and appropriate safeguards depending on the risk. Risks indicators include: sensitivity or highly personal nature of the data, vulnerability of the data subjects, the large scale of the processing activities, the systematic nature of the monitoring, the innovative use or application of technological solutions, the evaluation of individuals, the combination of datasets, the legal or similarly significant effect of automated decision making[137]. Professional ethical standards governing a particular research project would also be considered a safeguard. Where researchers deploy Artificial Intelligence systems, there is the same need for safeguards and oversight. IT and engineering research often lack ethical oversight, so the European Commission committed to embed 'ethics by design' for all future EU-funded AI projects[138].

Appropriate safeguards could include conducting a data protection impact assessment of likely risks for rights and freedoms of natural persons[139], appointing a data protection officer (mandatory in case of a public authority or body, regular and systematic monitoring of data subjects on a large scale, processing on a large scale of special categories of data)[140], notifying a data breach, without undue delay and no later than 72 hours, when the breach is likely to pose a risk to the rights and freedoms of the data subjects[141], guaranteeing data security[142], and data minimisation through pseudonymisation or (unless it would impair the research) anonymisation.

7. Recommendations

Dialogue between the scientific research community and data protection authorities is deepening, especially in the area of medical research. This needs to be intensified at EU level, because of the discrepancies across the Member States, misunderstanding about the GDPR, and concentration of potentially valuable social research data in a few private companies. As

well as inviting comments on this Preliminary Opinion, the EDPS would like to contribute to this dialogue by suggesting a number of areas for further work.

7.1 DPAs and ethical review boards

Data protection authorities and data protection officers increasingly engage with ethical questions in the development and deployment of digital technologies. They should engage more closely with ethical review boards. Genetic research in particular has implications not only for the subject of the DNA tests but others in his or her family or with shared characteristics in this and future generations[143]. Independent ethical committees could support the understanding of which activities qualify as genuine research and define the ethical standards referred to in the GDPR. Ethics committees can play a meaningful role in ensuring that the respect of human rights, including right to data protection, is embedded in the research project from the early planning stage. They are likely to continue to play an important role in ensuring that research projects are designed from the start with data protection principles in mind.

7.2 EU Codes of Conduct and certification for research activities

The GDPR requires Member States, supervisory authorities, the EDPB and the Commission to encourage the drawing up of codes of conduct to contribute to the proper application of the Regulation[144]. In research, codes of conduct can improve convergence of practices and increase confidence in compliance[145]. To achieve sufficient levels of harmonisation, codes of conduct at EU rather than national level may be preferable. They would also be beneficial for the free movement of researchers, a key aim of the European Research Area[146].

Although not covering data protection, a similar EU-wide project was achieved successfully in the past with the European Code of Conduct for Research Integrity[147] which contributed to harmonisation. Specialised codes might be particularly relevant for fields such as biobanking, genomic research or social networks research[148].

In addition, accredited certification bodies will be able to issue certifications to controllers or processors, such as data protection seals and marks, for a maximum and renewable period of three years. Their purpose is to demonstrate the compliance of processing operations with the Regulation[149].

Such codes of conduct and certifications under the GDPR could usefully address:

- requirements of valid consent as a lawful basis for data processing and/or safeguard;
- regime for special categories of personal data;
- legitimate interests pursued by researchers;
- pseudonymisation of research data and scientific publications;
- exercise of the rights of the data subjects in the context of the potential limitations of those rights;
- implementation of data protection by design in the field of research;
- transfers of personal data to third countries or international organisations[150];

- provision by private companies, particularly tech platforms, of data to independent researchers for specific projects, such as examining online manipulation and the dissemination of misinformation.

The research community is by nature heterogeneous, so a multiplicity of such instruments may be envisaged[151].

7.3 EU research framework programmes and data protection standards

Building on the considerable harmonisation efforts of the European Commission in the research area with Horizon 2020 and Horizon Europe, the next European Research and Innovation framework programme, can also support convergence across the Member States. Research projects seeking funding under the framework need to go through a vigorous ethics review process which starts with a self-assessment. It requires researchers to reflect on the design of the research project and prescribes that data protection requirements should be part of the ethical review process[152]. This process has considerable potential to align the data protection practices of all research institutions or independent researchers seeking to obtain European research funding. Researchers should seek guidance from data protection experts and authorities in the development of these research proposals[153].

7.4 Debate on a public interest ground for scientific research

Sharing personal data always involves a degree of risk to the people concerned by the data, including where the purpose is scientific research. There should be no loophole to the protection of fundamental rights, and uncertainty around what counts as 'scientific research' itself poses a risk of such loopholes emerging.

At the same time, there is growing concern about how digitisation has involved the exponential growth in data generation while also concentrating the control of the means for converting that data into valuable knowledge in the hands of a few powerful private companies. As explained above (section 2.4), some argue that it is undemocratic for these companies to monopolise such opportunities while shielding themselves from the scrutiny and accountability which independent researchers could provide, were only they able to examine the way information is disseminated by these platforms. Through their existing terms of service and privacy policies these companies allow themselves ample scope to determine how they wish to process personal data and with whom to share it. There therefore seems little obstacle to these terms of service providing for sharing data with genuine researchers operating within proper ethical governance. As this is not happening, there have been recent calls for regulated access across the EU to privately-held personal data for research purposes that serve a public interest, such as improving healthcare provision and addressing the climate crisis[154]. A public interest basis under data protection law for dominant companies to disclose data to researchers would need to be clearly formulated and laid down in EU or Member State law, as well as being accompanied by a rigorous proportionality test and appropriate safeguards against misuse and unlawful access.

The EDPS can help facilitate a debate on this matter with civil liberties groups, the research community and the major tech companies.

8. Conclusion

This Preliminary Opinion has sought to highlight the main challenges in the application of the GDPR to scientific research. Digitisation has created new potential for individual empowerment and addressing acute social issues like public health. It has also resulted in enormous accumulation and concentration personal data for the private benefit of a few global

companies, with the blurring of the boundaries between public interest, academic freedom and private gain[155]. Digital technology has been used to experiment on people online for private benefit. Data protection rules aim to ensure safety and transparency while minimising interference with ethical research that aim at generalisable knowledge and societal good. The GDPR serves in part to ensure accountability for such practices. There is no evidence that the GDPR itself hampers genuine scientific research. DPAs, ethics committees and the research community generally have a common interest in working together to help the advancement of knowledge, while ensuring people are not treated as mere data sets[156]. The EDPS in particular will be vigilant, mindful of how more time is needed to see how the special regime for data protection in the field of scientific research plays out on the ground.

NOTES

[1] OJ L 119, 4.5.2016, p. 1 (hereinafter "GDPR").
[2] OJ L 295, 21.11.2018, p. 39.
[3] OJ L 119, 4.5.2016, p. 89.
[4] See for instance Political Guidelines for the Next European Commission2019-2024; https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf
[5] The most complete review of perceptions and reality was conducted by the Health Ethics and Policy Lab, ETH, Zurich and published in July 2019 by DG EPRS of the European Parliament, https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2019)634447 (24 July 2019); see also MLC Foundation, Med law consult, 'GDPR and research' (16 November 2017). Available at https://www.medlaw.nl/nieuws/gdpr-and-research/ [this and all the articles in this opinion were accessed on 25.11.2019].
[6] There is some discussion as to whether the GDPR poses a barrier to research. Timmers et al refer to 'an inherent tension between critical care research and data protection'. Rabesandratana claims the GDPR is a 'serious impediment' for some researchers. Rumbold and Pierscionek on the other hand consider the GDPR 'will make little impact on biomedical data research', and 'will hopefully ensure that subjects continue to have trust in the integrity of their health care data and the medical research community' and 'will facilitate medical research, except where it is research not considered in the public interests.' Forgó does not foresee 'significant impact on ICT-supported medical research'. The Council of Canadian Academies in 2015 foresaw that 'opt-in by default', what was considered to be the GDPR standard, 'could have devastating consequences for researcher access to and use of data generated in health contexts'. See Adrian Thorogood, 'Canada: will privacy rules continue to favour open science?' (16 July 2018), 137(8) *Human genetics* 595; Kart Pormeister, 'Genetic data and the research exemption: is the GDPR going too far?' (4 May 2017), 7(2) *International Data Privacy Law* 137; Timmers et al, 'Will the EU Data Protection Regulation 2016/679 inhibit critical care research?' (2018), 27(1) *Medical Law Review* 59; John Rumbold and Barbara Pierscionek, 'A critique of the regulation of data science in healthcare research in the European Union' (2017), 18(1) *BMC Medical* Ethic; Nikolaus Forgó, 'My health data--your research: some preliminary thoughts on different values in the General Data Protection Regulation' (2014), 5(1) *International Data privacy Law* 54; Gauthier Chassang, 'The Impact of EU general data protection regulation on scientific research' (2017); Tania Rabesandratana, 'Researchers sound alarm on European data law', *Science*, 22 November 2019. 709 *ecancermedicalscience*. A report by German data protection authorities into the implementation asserts that 'the scope of the privileges accorded to science and scientific research in point (b) of Article 5(1) read in conjunction with Article 6(4) of the GDPR has proved to be too wide in practice; Independent German Federal and State Data Protection Supervisory Authorities, Report on Experience Gained in the Implementation of the GDPR, November 2019.
[7] Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC (Text with EEA relevance.) Article 13 of Regulation 2018/1725 is identical to GDPR Article 89(1).
[8] See Menno Mostert et al, 'Big Data in medical research and EU data protection law: challenges to the consent or anonymise approach' (2016), 24(7) *European Journal of Human Genetics* 956.
[9] For example, Recommendation CM/Rec(2019)2 of the Committee of Ministers to Member States on the protection of health-related data, 27 March 2019, states that scientific research requires consent or law authorising the use of the data.
[10] Speech by Giovanni Buttarelli (EDPS), 'The Impact of the General Data Protection Regulation on collaborative science in Europe and the European Cloud Initiative', in seminar organised by ISC Intelligence in Science, Brussels, 18 October 2016. Available at https://edps.europa.eu/press-publications/press-news/videos/impact-gdpr-collaborative-science_en.
[11] Edward S. Dove et al, 'Ethics review for international data-intensive research' (20 April 2016), 351 (6280) *Science* 1399; Deborah Mascalzoni et al, 'International Charter of Principles for Sharing Bio-Specimens and Data' (24 September 2014), 23 *European Journal of Human Genetics* 721.
[12] Gary King and Nathaniel Persily, A New Model for Industry-Academic Partnerships, American Political Science Association, 2019.
[13] See quote from Gary King Director of the Institute for Quantitative Social Science at Harvard University in Visvak, 'After scandal, Facebook restricts access to data-and cuts off academics' (9 May 2018), *The Wire*. Available at https://thewire.in/the-sciences/after-scandal-facebook-restricts-access-to-data-and-cuts-off-academics.
[14] Andrew G Reece and Christopher M Danforth, 'Instagram photos reveal predictive markers of depression' (8 August 2017), 6 *EPJ Data Science* 2017 15.
[15] Commission Community framework for state aid for research and development and innovation, OJ C 323/01, 30.12.2006, valid from 2006 to 2014, included a definition of 'research organisation'; but the current framework

has no such definition. The old definition was 'an entity, such as university or research institute, irrespective of its legal status (organised under public or private law) or way of financing, whose primary goal is to conduct fundamental research, industrial research or experimental development and to disseminate their results by way of teaching, publication or technology transfer; all profits are reinvested in these activities, the dissemination of their results or teaching; undertakings that can exert influence upon such an entity, in the quality of, for example, shareholders or members, shall enjoy no preferential access to the research capacities of such an entity or to the research results generated by it.'; Communication from the Commission, Framework for State aid for research and development and innovation, OJ C 198, 27.6.2014.

[16] Thomas J. May, 'Sociogenetic Risks — Ancestry DNA Testing, Third-Party Identity, and Protection of Privacy' (July 2018), 379(5) *The New England Journal of Medicine* 410.

[17] Molly McCluskey, 'Public Universities Get an Education in Private Industry' (3 April 2017), *The Atlantic*. Available at https://www.theatlantic.com/education/archive/2017/04/public-universities-get-an-education-in-private-industry/521379/.

[18] Google funded 100 public research papers since 2009 but most did not disclose; Brody Mullins and Jack Nicas, 'Paying Professors: Inside Google's Academic Influence Campaign' (14 July 2017), *The Wall Street Journal*. Available at https://www.wsj.com/articles/paying-professors-inside-googles-academic-influence-campaign-1499785286.

[19] Lydia DePillis, 'Amazon gets an edge with its secret squad of PhD economists' (13 March 2019), *CNN*. Available at https://edition.cnn.com/2019/03/13/tech/amazon-economists/index.html.

[20] Cara McGoogan, 'NHS illegally handed Google firm 1.6m patient records, UK data watchdog finds' (3 July 2017), *Telegraph*. Available at https://www.telegraph.co.uk/technology/2017/07/03/googles-deepmind-nhs-misused-patient-data-trial-watchdog-says/.

[21] Ibid.

[22] Information Commissioner's Office (ICO), Report to the Parliament of 6 November 2018, 'Investigation into the use of data analytics in political campaigns', pp. 55-58.

[23] Namely matured in the context of the Psychometric Centre at the Cambridge University with the afterwards developed app "My personality".

[24] ICO Report, op.cit, 'Investigation into the use of data analytics in political campaigns', p. 38.

[25] Kalokairinou, L., Howard, H.C., Slokenberga, S. et al., Legislation of direct-to-consumer genetic testing in Europe: a fragmented regulatory landscape, J Community Genet (2018) 9: 117, https://doi.org/10.1007/s12687-017-0344-2

[26] Henri-Corto Stoekle, Marie-France Mamzer-Bruneel, Guillaume Vogt and Christian Hervé, '23andMe: a new two-sided data-banking market model' (2016), 17 *BMC Medical Ethics* 19.

[27] Megan Molteni, 'These DNA Startups want to put all of you on the blockchain' (16 November 2018), *Wired*. Available at https://www.wired.com/story/these-dna-startups-want-to-put-all-of-you-on-the-blockchain/.

[28] See the Personal Genome Project. Available at https://www.personalgenomes.org/.

[29] See 'The Future of Informed Consent in Research and Translational Medicine: A Century of Law, Ethics and Innovation' (March 2018), 46(1) *The Journal of Law, Medicine and Ethics*, in particular contributions of Effy Vayena and Alessandro Blassime, 'Health Research with Big Data: Time for Systematic Oversight' pp. 119-129, and Susan M. Wolf et al, 'The Past Present and Future of Informed Consent in Research and Translational Medicine, pp. 7-11.

[30] Chavie Lieber, 'Tech companies use "persuasive design" to get us hooked. Psychologists say it's unethical' (8 August 2018) *Vox*. Available at https://www.vox.com/2018/8/8/17664580/persuasive-technology-psychology.

[31] Josh Constine, 'Facebook pays teens to install VPN that spies on them' (30 January 2019), *Techcrunch*. Available at https://techcrunch.com/2019/01/29/facebook-project-atlas/.

[32] Adam D. I. Kramer, Jamie E. Guillory and Jeffrey T. Hancock, 'Experimental evidence of massive-scale emotional contagion through social networks' (June 2014) National Academy of Sciences. Available at: https://www.pnas.org/content/111/24/8788.

[33] 'OKCupid experiments with 'bad' dating matches' (29 July 2014), *BBC NEWS Technology*. Available at https://www.bbc.com/news/technology-28542642.

[34] See https://www.wired.com/2016/05/okcupid-study-reveals-perils-big-data-science/

[35] Shannon Liao, 'Amazon put unwanted sponsored products in customers' baby registries' (28 November 2018), *The Verge*. Available at https://www.theverge.com/2018/11/28/18116274/amazon-baby-registry-unwanted-sponsored-products-ads.

[36] See Stanford Persuasive Technology Lab. Available at http://captology.stanford.edu/.

[37] Meredith Whittaker et al, 'AI Now Report 2018' (December 2018), AI NOW Institute. Available at https://ainowinstitute.org/AI_Now_2018_Report.pdf.

[38] An example of such research is Jonathan Albright's study into the spread of anti-Semitism and propaganda via social media platforms; see https://medium.com/s/the-micro-propaganda-machine/. Social Science One is an attempted collaboration between academics and private industry to analyse the latter's troves of information for wider societal benefit; https://socialscience.one/ .

[39] See for example https://www.theguardian.com/technology/2018/mar/16/silicon-valley-internal-work-spying-surveillance-leakers ; https://www.businessinsider.com/secretive-tech-companies-apple-google-palantir-2019-9?r=US&IR=T ; and https://www.ipwatchdog.com/2019/09/25/dark-side-secrecy-theranos-can-teach-us-trade-secrets-regulation-innovation/id=113907/ on the lessons from the Theranos scandal; Natasha Lomas, 'Europe's parliament calls for full audit of Facebook in wake of breach scandal' (25 October 2018), *Techcrunch*. Available at https://techcrunch.com/2018/10/25/europes-parliament-calls-for-full-audit-of-facebook-in-wake-of-breach-scandal/.

[40] Natasha Lomas, 'Facebook accused of blocking wider efforts to study its ad platform' (29 April 2019), *TechCrunch*. Available at https://techcrunch.com/2019/04/29/facebook-accused-of-blocking-wider-efforts-to-study-its-ad-platform/; Jeremy B. Merrill, 'Facebook moves to block Ad Transparency Tools- including ours' (28 January 2019), *ProPublica*. Available at https://www.propublica.org/article/facebook-blocks-ad-transparency-tools.

[41] Letter to Google, Facebook and Twitter from ten researchers and co-signed by over 60 academics, 27.3.2019 refers to research into 'inauthentic accounts and behaviour on social media, political and issue-based advertising, microtargeting and ad placements by political parties and other entities, the effectiveness of self-regulation measures to counter disinformation. The European Commission urged Facebook to agree with the European research community a 'framework that facilitates access to a broader set of data, in compliance with privacy rules, in order to allow a comprehensive assessment of the Code of Practice implementation'; European Commission, Code of Practice on Disinformation: Intermediate Targeted Monitoring – March Reports.

[42] https://socialscience.one/blog/public-statement-european-advisory-committee-social-science-one

[43] See the 2018 report on privacy policies by the Norwegian Consumer Council; https://www.forbrukerradet.no/side/250000-words-of-app-terms-and-conditions/

[44] Commission Staff Working Document, Guidance on sharing private sector data in the European data economy Accompanying the document Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, "Towards a common European data space" Brussels, 25.4.2018 SWD(2018) 125 final; See also Begoña Gonzalez Otero, , Evaluating the EC Private Data Sharing Principles: Setting a Mantra for Artificial Intelligence Nirvana? (November 28, 2018); https://ssrn.com/abstract=3382330

[45] The aim of the advancement of knowledge is the core aspect of pure basic research as understood by the OECD, Glossary of Statistical Terms, available at: https://stats.oecd.org/glossary/detail.asp?ID=2206; Lisa Bortolotti and Bert Heinrichs, 'Delimiting the concept of research: an ethical perspective' (June 2007), 28(3) *Theoretical Medicine and Bioethics* 157.

[46] European Commission, Open innovation, open science, open to the world: A vision for Europe, 30 May 2016.

[47] *The American Heritage: Medical Dictionary* (2007) Boston: Houghton Mifflin Co.

[48] Forgó, op. cit.

[49] Compare however with the exclusion of for-profit undertakings from the exceptions to research institutions according to the European Commission's Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market COM(2016) 593, final Recital 11: "Research organisations across the Union encompass a wide variety of entities the primary goal of which is to conduct scientific research or to do so together with the provision of educational services. Due to the diversity of such entities, it is important to have a common understanding of the beneficiaries of the exception. Despite different legal forms and structures, research organisations across Member States generally have in common that they act either on a not for profit basis or in the context of a public-interest mission recognised by the State. Such a public-interest mission may, for example, be reflected through public funding or through provisions in national laws or public contracts. At the same time, organisations upon which commercial undertakings have a decisive influence allowing them to exercise control because of structural situations such as their quality of shareholders or members, which may result in preferential access to the results of the research, should not be considered research organisations for the purposes of this Directive."

[50] *Sorguç v. Turkey* App no 17089/03 (ECHR, 23 June 2009), par. 35.

[51] Article 85 GDPR. This provision was absent under the previous Directive 95/46.

[52] Article 85(2) of GDPR provides that in order to reconcile the right to the protection of personal data with the right to freedom of expression, including processing for journalistic purposes and the purposes of academic, artistic or literary expression, Member States shall provide exemptions or derogations from Chapter II (principles), Chapter III (rights of the data subject), Chapter IV (controller and processor), Chapter V (transfer of personal data to third countries or international organisations), Chapter VI (independent supervisory authorities), Chapter VII (cooperation and consistency) and Chapter IX (specific data processing situations).

[53] Proposal for a Directive of the European Parliament and of the Council on copyright in the Digital Single Market, COM(2016)593 final, art 3.

[54] Ibid, art 2.

[55] Recital 12 of GDPR.

[56] Recital 157 of GDPR.; See US Federal 'Common Rule' which defines research as 'a systematic investigation… designed to develop or contribute to generalizable knowledge.'; US Department of Health and Human Services,

ORI The Office of Research Integrity, Definitions, ORI Introduction to RCR: Chapter 3 'The Protection of Human Rights'. Available at https://ori.hhs.gov/content/chapter-3-The-Protection-of-Human-Subjects-Definitions.

[57] Recital 159 of GDPR.

[58] Ibid; See the Opinion of AG Mancini in C-234/83 *Gesamthochschule Duisburg v Hauptzollamt München-Mitte* [1985] on interpretation of 'scientific activities' in the context of the legislation relating to custom duties (first indent of Article 3(2) of Regulation No 1798/75): 'scientific activities must be interpreted as including activities carried on by a public or private establishment engaged in education or research for the purpose of further the acquisition, development, exposition or dissemination of scientific knowledge (…)'.

[59] Speech by Giovanni Buttarelli (EDPS) during the Fifth World Congress for Freedom of Scientific research, 12 April 2018. Available at https://edps.europa.eu/sites/edp/files/publication/18-04-12_fifth_world_congress_freedom_scientific_research_en.pdf, p. 2.

[60] Article 29 Working Party, Guidelines on consent under Regulation 2016/679 (wp259rev.01) (as last revised and adopted on 10 April 2018), pp. 27-30.

[61] Speech by Giovanni Buttarelli (12 April 2018), op. cit., p. 2.

[62] See the Opinion of AG Mancini in Case 234/83 *Gesamthochschule Duisburg v Hauptzollamt München-Mitte* [1985] on interpretation of 'scientific activities' in the context of the legislation relating to custom duties (first indent of Article 3(2) of Regulation No 1798/75): 'scientific activities must be interpreted as including activities carried on by a public or private establishment engaged in education or research for the purpose of further the acquisition, development, exposition or dissemination of scientific knowledge (…)'.

[63] All projects receiving Horizon 2020 funding are required to make sure that any peer-reviewed journal article they publish is openly accessible, free of charge, see http://ec.europa.eu/research/participants/docs/h2020-funding-guide/grants/grant-management/dissemination-of-results_en.htm.

[64] More information can be found on the OpenAIRE website: https://www.openaire.eu/.

[65] European Commission, 'A Digital Single Market Strategy for Europe' COM(2015) 192 final, 15; European Commission's press release, 'European Cloud Initiative to give Europe a global lead in the data-driven economy' (19 April 2016): "developing a European open science cloud for European researchers and their global scientific collaborators by integrating and consolidating e-infrastructure platforms, federating existing scientific clouds and research infrastructures, and supporting the development of cloud-based services".

[66] European Commission, Proposal for a Council Regulation on establishing the European High Performance Computing Joint Undertaking COM(2018) 8 final.

[67] More information can be found on the European Open Science Cloud for Research Pilot Project website: https://eoscpilot.eu.

[68] Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information, OJ L 175 (27 June 2013), pp. 1-8. The PSI Directive explicitly states that it affects in no way the level of protection of individuals with regard to the processing of personal data; Article 1(4) of PSI Directive.

[69] European Commission, Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information (recast) COM(2018) 234 final, Article 10.

[70] European Commission, Recommendation of 25 April 2018 on access to and preservation of scientific information C(2018) 2375 final, p. 5.

[71] For insights on those developments from the research community perspective, see for example: Katrin Schaar, 'What is important for data protection in science in the future? General and specific changes in data protection for scientific use resulting from the EU General Data Protection Regulation' (July 2016), 258 *RatSWD Working Paper Series*, p. 2; Eckard Kämper, 'Risiken sozialwissenschaftlicher Forschung? Forschungsethik, Datenschutz und Schutz von Persönlichkeitsrechten in den Sozial- und Verhaltenswissenschaften' (March 2016), 255 *RatSWD Working Paper*, p. 4; John M.M. Rumbold and Barbara K. Pierscionek, 'A critique of the Regulation of data science in healthcare research in the European Union' (8 April 2017), 18 *BMC Medical Ethics* 27.

[72] Evelyne Shuster, 'Fifty Years Later: The Significance of the Nuremberg Code' (13 November 1997), 337 *The New England Journal of Medicine* 1436. Available at https://www.nejm.org/doi/full/10.1056/NEJM199711133372006.

[73] Declaration of Helsinki-Ethical principles for medical research involving human subjects (first adopted 1964, amended 2013), paragraphs 24 and 32. Beyond those main instruments, there is the UNESCO Universal Declaration on Bioethics and Human Rights of 2005; the World Medical Association The Council for International Organisations of Medical Sciences (CIOMS), *International Ethical Guidelines for Biomedical Research Involving Human Subjects* (CIOMS, 2002); The Council for International Organisations of Medical Sciences (CIOMS), *International Ethical Guidelines for Epidemiological Studies* (CIOMS, 2009); World Medical Association (WMA) Declaration of Taipei on Ethical Considerations Regarding Health Databases and Biobanks (2002). The European Commission has also issued two helpful guidelines: Opinion Nr 17 on Ethical Aspects of Clinical Research in Developing Countries (2003), available at https://op.europa.eu/en/publication-detail/-/publication/6339dcbf-c156-4e7f-9e43-9928acf82118/language-en/format-PDF/source-77404483 and the Horizon 2020 Guidance on How to Complete your Ethics Self-Assessment (2019), available at

https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf.

[74] The states being California, New York and Maryland, Virginia; https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3954467/

[75] Council of Europe, Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine (Oviedo, 4 April 1997). Available at https://www.coe.int/en/web/conventions/full-list/-/conventions/rms/090000168007cf98.

[76] Ibid, Article 16.

[77] Declaration of Helsinki, paragraphs. 25 to 32.

[78] European Commission, European Textbook on Ethics in Research (2010).

[79] Michal Koščík and Matěj Myška, 'Data protection and codes of conduct in collaborative research' (15 January 2018), 32 *International Review of Law, Computers & Technology* 141. As an example, researchers who wish to collect personal data in Finland must fill in the description of the scientific research data file and they usually need to submit it to the ethical review boards. See the Consortium of European Social Science Data Archives' (CESSDA) pragmatic overview of national practices: https://www.cessda.eu/Research-Infrastructure/Training/Expert-Tour-Guide-on-Data-Management/5.-Protect/Processing-personal-data/Diversity-in-data-protection.

[80] See the Network's website for more information: http://www.eurecnet.org/index.html.

[81] European Commission, Ethics and data protection, 14 November 2018. Available at: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-data-protection_en.pdf.

[82] Menno Mostert et al, 'From Privacy to Data Protection in the EU: Implications for Big Data Health Research' (11 December 2017), 25(1) *European Journal of Health Law* 43, p. 44.

[83] European Medicines Agency Policy on Publication of Clinical Data for Medical Products for Human Use 0070 (21 March 2019); Government of Canada, Forward Regulatory Plan 2019-2021: Regulations amending the Food and Drug Regulations and Medical Devices Regulations-Recall of Therapeutic Products; U.S. Food and Drug Administration pilot program to evaluate the sharing of clinical trial documents.

[84] Timmers et al, op. cit., pp. 6-7.

[85] Mark Sheehan, 'Can broad consent be informed consent?' (November 2011), 4(3) *Public Health Ethics* 226; Graeme Laurie et al., 'A Review of evidence relating to harm resulting from uses of health and biomedical data' (30 June 2014), *Nuffield Council on Bioethics*.

[86] See for instance Eline M. Bunnik et al., 'A tiered-layered-staged model for informed consent in personal genome testing' (21 November 2012), 21 *European journal of human genetics* 596.

[87] See for example, Kristin Solum Steinsbekk et al., 'Broad consent versus dynamic consent in biobank research: Is passive participation an ethical problem?' (September 2013), 21(9) *European journal of human genetics* 897.

[88] Timmers et al, op. cit.

[89] See Articles 3(2)(b), (d), (e), 4(a) and 5 of Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use, OJ L 121 (01.05.2001). See also the Regulation (EU) 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use and repealing Directive 2001/20/EC, OJEU L 158 (27.05.2014), in particular Articles 28 to 34.

[90] Regulation (EU) 536/2014, op. cit., chapter V, Article 2(2)(21).

[91] Ibid, Article 29(1).

[92] Ibid, Recital 30.

[93] Ibid, Recital 31.

[94] Ibid, Recitals 27 and 29.

[95] Ibid, Recital 76 on withdrawal of consent indicates that "without prejudice to Directive 95/46/EC, the withdrawal of informed consent should not affect the results of activities already carried out, such as the storage and use of data obtained on the basis of informed consent before withdrawal."

[96] Ibid, Recital 29.

[97] European Data Protection Board, Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical Trials Regulation (CTR) and the General Data Protection regulation (GDPR), 23 January 2019.

[98] This Article provides that "any limitation on the exercise of the rights and freedoms recognised by this Charter must be provided for by law and respect the essence of those rights and freedoms. Subject to the principle of proportionality, limitations may be made only if they are necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of other".

[99] See for instance Recital 34 to 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 (replaced by the GDPR): "…Member States must also be authorized, when justified by grounds of important public interest, to derogate from the prohibition on processing sensitive categories of data where

important reasons of public interest so justify in areas such as public health and social protection - especially in order to ensure the quality and cost-effectiveness of the procedures used for settling claims for benefits and services in the health insurance system - scientific research and government statistics; whereas it is incumbent on them, however, to provide specific and suitable safeguards so as to protect the fundamental rights and the privacy of individuals".

[100] See Recital 33 GDPR. See also Oviedo Convention, Helsinki Declaration and Taipei. Also of interest is House Bill 917 of State of Maryland, April 2016, para. 13-2002: "A person may not conduct research using a human subject unless the person conducts the research in accordance with the federal regulations on the protection of human subjects".

[101] Recital 157 of GDPR.

[102] Directive 95/46/EC did not provide for a special regime but allowed Member States to adopt rules derogating from the general rules. See for example Kärt Pormeister, 'Genetic research and applicable law: the intra-EU conflict of laws as a regulatory challenge to cross-border genetic research' (10 November 2018), 5(3) *Journal of Law and the Biosciences* 706.

[103] See EDPS, Assessing the necessity of measures that limit the fundamental right to the protection of personal data: A Toolkit, 11 April 2017 and discussion (footnote 8) of CJEU judgments in *Schrems*, CJEU Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland* 8 April 2014 and *Tele2 Sverige AB*

[104] *Delcourt v Belgium* App no. 2689/65 (ECHR, 17 January 1970), and *Klass and others v Germany* App no. 5029/71 (ECHR, 6 September 1978); Case C-293/12 *Digital Rights Ireland* [2014] ECLI:EU:C:2014:238.

[105] Article 4(1)(11) of GDPR.

[106] Article 29 Working Party, Guidelines on consent under Regulation 2016/679, WP259, 28 November 2017.

[107] See C-673/17 *Planet49* judgment of 1 October 2019, and C-61/19 Orange Romania (pending).

[108] Article 29 Working Party, Guidelines on consent under Regulation 2016/679, op. cit., pp. 5-7.

[109] Recital 43 of GDPR.

[110] EDPB, Opinion 3/2019 concerning the Questions and Answers on the interplay between the CTR and the GDPR, op. cit.

[111] Article 29 Working Party, Guidelines on consent under Regulation 2016/679, op. cit., p. 11.

[112] See also Article 29 Working Party, Opinion 03/2013 on purpose limitation, WP203, 2 April 2013.

[113] See *Planet 49* para 52 and 54.

[114] Article 29 Working Party, Guidelines on consent under Regulation 2016/679, op. cit., p. 28.

[115] Ibid. See also the recent Guidance issued by the Association for German Supervisory Authorities on the interplay between recital 33 and the definition of consent in the GDPR (3 April 2019) confirms this approach. Available here in German: https://www.datenschutzkonferenz-online.de/media/dskb/20190405_auslegung_bestimmte_bereiche_wiss_forschung.pdf

[116] Article 29 Data Protection Working Party, WP 258 Opinion on some key issues of the Law Enforcement Directive (EU 2016/680), Adopted on 29 November 2017

[117] Article 7(3) of GDPR.

[118] Article 17(1)(b) and (3) of GDPR.

[119] Article 13 of GDPR. In a research context providing such information may sometimes prove problematic where the data was collected long time ago; for instance, where a university hospital intends to use data from decades old patients' records.

[120] Article 14 of GDPR.

[121] Article 29 Working Party, Guidelines on transparency under regulation 2016/679, WP260, adopted on 29 November 2017 and last revised on 11 April 2018, pp. 28-31.

[122] Articles 13(3) and 14(4) GDPR.

[123] European Commission, European Textbook on Ethics in Research, op. cit.

[124] See C-553/07 *College van burgemeester en wethouders van Rotterdam v. M.E.E. Rijkeboer* [2009], ECLI:EU:C:2009:293, par. 49-54.

[125] Article 14(5)(b) of GDPR.

[126] EDPS Opinion 10/2017 on safeguards and derogations under Article 89 GDPR in the context of a proposal for a Regulation on integrated farm statistics, 20 November 2017.

[127] Article 5(1)(b) of GDPR.

[128] Article 29 Working Party, Opinion 03/2013 on purpose limitation, op. cit., pp. 23-28.

[129] See Information Commissioner's Office, Investigation into the use of data analytics in political campaigns, Investigation update. Available at: https://ico.org.uk/media/action-weve-taken/2259371/investigation-into-data-analytics-for-political-purposes-update.pdf.

[130] Article 6(1)(b) of Directive 95/46/EC, op. cit.

[131] Article 29 Working Party, Opinion 03/2013 on purpose limitation, op. cit., p.28

[132] Article 8(2) of the EU Charter of Fundamental Rights provides: "Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law" (emphasis added).

[133] Ibid, pp. 12, 19.

[134] Timmers et al, op. cit. p. 15 and in particular p. 18: 'The health care data are fundamentally different from posts on social media. They are the professional result of health care and based on the best possible understanding and scientific data of the patient's condition. Outside critical care they will also reflect the choices the patient has made about possible diagnostic and therapeutic interventions but win those limitations as set by the patient, the data should reflect the professional norm comparable to similar patients.'

[135] *Handyside vs the UK* App no. 5493/72 (ECHR, 7 December 1976); *Leander v. Sweden* App no. 9248/81 (ECHR, 26.03.1987); see also EDPS, Assessing the necessity of measures that limit the fundamental right to the protection of personal data: A Toolkit, op. cit.

[136] See recommendations 16-23 of the Berlin Data Ethics Commission, Opinion of the Data Ethics Commission-Executive Summary, 22 October 2019. Available at https://www.bmjv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN.html;jsessionid=1B71C1E6D363C833EC7F485C2AF205AD.1_cid297?nn=11678512.

[137] Article 29 Working Party, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 4 April 2017, p. 11.

[138] https://ec.europa.eu/transparency/regdoc/rep/1/2018/EN/COM-2018-237-F1-EN-MAIN-PART-1.PDF

[139] Article 35 of GDPR.

[140] Article 37 of GDPR.

[141] Articles 33 and 34 of GDPR.

[142] Article 32 of GDPR.

[143] See Kärt Pormeister, 'Genetic Data and the research exemption: is the GDPR going too far?' (4 May 2017) 7(2) *International Data Privacy Law* 137.

[144] Article 40(1) of GDPR.

[145] They could rely on numerous documents such as the OECD Recommendation of the Council on Health Data Governance, 13 December 2016, available at https://www.oecd.org/health/health-systems/Recommendation-of-OECD-Council-on-Health-Data-Governance-Booklet.pdf; eTRIKS, Code of practice on secondary use of medical data in scientific research projects, 27 August 2014, available at https://www.etriks.org/wp-content/uploads/2014/12/Code-of-Practice-on-Secondary-Use-of-Medical-Data-with-recognition.pdf); EU Cloud Code of Conduct, EU Data Protection Code of Conduct for cloud service providers, March 2019, available at https://eucoc.cloud/fileadmin/cloud-coc/files/European_Cloud_Code_of_Conduct.pdf; Article 29 Working Party, Review of the Commission's code om mobile health applications, 10 April 2017, available at https://ec.europa.eu/newsroom/document.cfm?doc_id=44371 ; Italian Code of conduct and professional practice applying to processing of personal data for statistical and scientific purposes (https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/1115480).

[146] Koščík and Myška, op. cit. p. 2.

[147] Drafted by ALLEA, the code was last revised in 2017. The text is available at https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020-ethics_code-of-conduct_en.pdf.

[148] Koščík and Myška, op. cit., p. 10. Two pan-European drafting initiatives are under way. First, a Code of Conduct for Health Research is currently being drafted by the Biobanking and Biomolecular Resources Research Infrastructure-European Research Infrastructure Consortium (BBMRI-ERIC). The aim is a comprehensive sector-specific code. In the context of clinical trials, it will focus on the secondary use of data which are not covered by the Clinical Trials Directive. Second, the initiative led by GÉANT, the pan-European data network for the research and education community. The Code of Conduct relates to processing of personal data for online access management purposes in the research and education sector.

[149] Articles 42(1), (5), (7) and 43 GDPR. See for example, the data seals in the United Kingdom, available at https://ico.org.uk/for-organisations/improve-your- practices/privacy-seals/.

[150] See Article 40(2)(b), (c), (d), (h) and (j) of GDPR.

[151] Koščík and Myška, op. cit., p. 10. This paper suggests the Association of All European Academies (ALLEA) and the European University Association (EUA).

[152] The H2020 Self-assessment guidelines are available at: http://ec.europa.eu/research/participants/portal/doc/call/h2020/h2020-msca-if-2015/1645175-h2020_-_guidance_ethics_self_assess_en.pdf. For more information on the process, the online manual on the procedure is available at: https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/ethics_en.htm.

[153] See European Commission advice on ethics and data protection to applicants for research funding https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-data-protection_en.pdf

[154] See Berlin Data Ethics Commission (footnote 136). An analogy may be drawn with the field of patent law for instance, where the EU Regulation on compulsory licensing of patents for the manufacture of pharmaceutical products for export to countries with public health problems outside the EU, provides for access to the patent information against a fee, or in the case of law enforcement and national security.

[155] United Nations Conference on Trade and Development, Digital Economy Report 2019, September 2019; https://unctad.org/en/PublicationsLibrary/der2019_en.pdf?user=46

[156] Os Keyes, 'The Gardener's vision of data' (6 May 2019) *Real Life*. Available at https://reallifemag.com/the-gardeners-vision-of-data/.