Genetic Data and the Data Protection Regulation: A good start, but problems remain

Dara Hallinan, Fraunhofer Institute for Systems and Innovation Research
Michael Friedewald, Fraunhofer Institute for Systems and Innovation Research
Paul de Hert, Vrije Universiteit Brussel
**Genetic Data and the Data Protection Regulation: A good start, but problems remain**

Does the approach to genetic data in the proposed Data Protection Regulation resolve the problems with the standard of fundamental rights protection for genetic data in the Data Protection Directive?

Dara Hallinan, Michael Friedewald; Fraunhofer Institute for Systems and Innovation Research
Paul De Hert; Vrije Universiteit Brussel

**Abstract:** Genes are the recipe, which is followed to create each person. Genetic data is data which refers to these genes. As genetic data is a representation of the biological architecture of the individual, it has been seen to have a number of unique features. Data protection law deals with the processing of ‘personal data’ and is a key area of law in the regulation of the use of genetic data. It is currently elaborated at European level by Directive 95/46. Owing to the unique qualities of genetic data, there have been numerous criticisms of the Directive’s suitability to achieve its legislative goals in relation to genetic data. One of these goals was to protect the fundamental rights of the data subject threatened through data processing. It has been suggested that the Directive did not recognise the sensitivity of genetic data and that it ignored a number of legitimate interests in this data (in particular interests which multiple data subjects may have and those which may remain in anonymous data). In 2012, the first results of a reform process of EU data protection law were released. These results included a draft Regulation (to replace the Directive) which outlined a novel approach to genetic data. Given that the goals of the Regulation have not changed from those of the Directive, this article considers whether the approach to genetic data in the Regulation addresses the questions posed by genetic data to the suitability of the Data Protection Directive to protect individuals’ fundamental rights. It concludes that the Regulation has rectified the lack of recognition of sensitivity, but still stutters in recognising a number of legitimate interests.

**Keywords:** Data Protection, Data Protection Directive, Data Protection Regulation, Data Protection Reform, Genetic, Genomic, Genetic Data, DNA, Bioethics

**Introduction**

Genes are the recipe, which is followed to create each person. Genetic data is data which refers to these genes. This can include data drawn from direct analysis of these

---

genes (DNA analysis), data drawn from observation of what a specific set of genes creates (phenotype analysis – people with blue eyes have...genetic features) or even social conclusions associated with the observation of a particular genetic quality (people with blue eyes are bad swimmers). The possibilities offered by genetic data to directly engage with the biological makeup of the human are unique and numerous. At their best they offer fresh insight and approaches to medicine, at their worst they are used as the foundations for horrific acts of discrimination.

Data protection law deals with the processing of ‘personal data’. It is currently elaborated at European level by Directive 95/46. One of the key aims of this Directive was to protect the fundamental rights of the data subject threatened through the processing of data. Whilst genetic data are not specifically mentioned in the Directive, they fulfil the criteria of ‘personal data’ and thus fall within its scope. However, the efficacy of a legislative instrument depends on how well it fits the substance and context of regulation. In the case of the Data Protection Directive, the substance of regulation was ‘data’. It has consistently been pointed out that genetic data display novel characteristics in relation to ‘normal’ data – the data around which the Directive was drafted.

In 2009, the Commission began considering a reform of the data protection framework. This process is expected to take the next few years at least, and will doubtless undergo numerous changes. However, in 2012, the Commission released the first tangible results, including a General Data Protection Regulation, intended to replace the Data Protection Directive. The Regulation retains the legislative aims of the Directive. However, in dealing with genetic data under the Regulation, the Commission decided to take an innovative approach. Genetic data are now clearly defined and classified as ‘sensitive’ under Article 9.

Given that the goals of the Regulation have not changed from those of the Directive, this article considers whether the choices made by the Commission in the regulation of genetic data address the questions posed by genetic data to the suitability of the Directive to protect individuals’ fundamental rights.

The Article begins, by clarifying the themes under discussion. First, a brief explanation of genetic data is given (part 1). Then, an explanation of the general construction of data protection is given, describing the forms of data it recognises – anonymous, normal and sensitive – and describing the three tiered standard of protection it employs (part 2). Finally, the article enters into a more specific account of the difficulties posed to data protection by genetic data. The article clarifies, in particular, the specific attributes of

---

2 The reform proposal also included the Police and Criminal Justice Data Protection Directive intended to replace the Framework Decision 2008/977/JHA. This article will only focus on the draft Data Protection Regulation.

3 This is thus an analysis aimed primarily at considering whether the proposals respond to the original problems with the Directive. There are a further set of problems with the content of the Commission’s approach in the proposed Regulation – for example relating to the potential breadth of the definition of genetic data. This second set of problems is only briefly touched upon in this paper.
genetic data as opposed to ‘normal data’. Accordingly, the article explains how genetic data thus represent a change in the ‘substance’ of regulation and therefore pose questions as to whether the Directive is a suitable tool in ensuring the protection of fundamental rights. Two forms of problem are highlighted. First, that the protection offered by the Directive did not reflect the sensitivity of genetic data and was insubstantial. Second, that there are a broader range of fundamental rights interests relevant in genetic data than in normal personal data, which the structure of the Directive excludes and thus ignores. (part 3).

This is followed by an elaboration of the first set of problems: The insubstantiality of the protection offered to genetic data by the framework. The Directive perceives that the processing of certain forms of data can pose a greater threat to fundamental rights. Yet, genetic data are not mentioned in the Directive, meaning they do not qualify for the higher standard of protection offered to ‘sensitive’ data. In principle, the Directive does not distinguish their treatment from that of other, ‘normal data’. However, other legal sources suggest genetic data have qualities which make their collection and processing particularly sensitive. This lack of recognition demonstrates somewhat of a misalignment of the construction of the data protection framework with the recognised ‘sensitivity’ of genetic data (part 4).

The article then elaborates the second set of problems. The Directive constructs a framework aimed at providing procedural justice for data subjects. For procedural justice to be legitimate and effective, the framework needs to be capable of recognizing all the rights and interests which could be affected, and all the parties to whom these rights belong, in any given processing situation. However, the Directive employs concepts such as ‘personal data’, ‘identifiability’ and ‘anonymity’ as gateways to accessing its protection. These are concepts which shut out the recognition of legitimate interests held in genetic data – in particular interests which multiple data subjects may have and those which may remain in anonymous data (part 5).

A brief explanation of the approach to genetic data, as imagined in the proposed data protection Regulation then follows. The Regulation’s approach to genetic data is considered, explaining that a definition has been offered and genetic data have been classified as ‘sensitive’ under Article 9. As a consequence of this classification, genetic data qualifies for a higher standard of protection than ‘normal data’. This classification limits the grounds under which processors can legitimately process data, puts extra requirements on processors to make sure their processing is legitimate and interests in data are considered, and provides mechanisms for the external oversight of processing and punishment of illegitimate processing (part 6).

Finally, the article considers whether the issues identified with the Directive’s engagement with genetic data are solved by the approach of the proposed Regulation. It is observed that the classification of genetic data as ‘sensitive’ under Article 9 addresses the lack of recognition of the sensitivity of genetic data in the Directive. This is a step that will strengthen the data subject’s fundamental rights without placing unjustifiable prohibitions on the processing of genetic data (part 7).

However, the proposed Regulation has done very little toward addressing the interests in data which lie beyond the scope of protection offered by the Directive, but which are present in genetic data. The Regulation does not address the possibility of multiple data subjects, and does not address the problems in the concept of anonymity (although there
may be some hope in an ambiguous interpretation of the scope of protection extended by Article 9) (part 8).

Overall, the article finds that the approach of the Regulation is a step in the right direction but that certain problematic issues remain unaddressed. The decision to define genetic data, and its classification as sensitive, address the Directive’s silence and provide a framework offering an increased standard of protection for the individual, which simultaneously maintains legislative and practical flexibility in the processing of genetic data. Unfortunately, the continuity in concepts and principles in the Regulation still leaves little space to recognize a number of legitimate interests in genetic data.

1. **What is genetic data?**

Our genes are the complete set of biological instructions which define what an organism is like, physically, and also, in part, socially. Broadly defined, genetic data can be understood to include any information relating to genes. There are two types of genetic data: 1: Information about what a specific genetic architecture is (the genetic code for eye colour in one individual – genotype). 2: What a specific architecture signifies (the resulting eye colour – phenotype). Architecture implies significance, while significance can also infer knowledge about architecture (people with a certain eye colour have a specific architecture, people with a specific architecture have a certain eye colour).

Traits can vary broadly in form, describing not only physical attributes, but also the current, or future, medical or social status of an individual (and potentially groups of individuals). Each piece of genetic information may be linked to, and thus used as an indicator for, the presence of further specific forms of architecture, or traits (a person with trait X is also likely to have trait Y).

Finally, genetic data gains significance as it is interpreted in a social context to make social evaluations (e.g. people with a certain eye colour are bad swimmers).

On the one hand, each individual’s specific genetic make-up is unique. On the other hand, specific architectural features will rarely be unique to one individual. Accordingly, if a feature of architecture has significance, it can often be generalised to infer knowledge about others who may be assumed to share that architecture (for example blood relatives – including ancestors and future progeny) but a number of other shared biological groups could be imagined. Genetic data may thus also have relevance for successive generations. (Taylor 2012, pp.41-62)

The two forms of genetic information represent two sides of the puzzle of the ‘nature’ (as opposed to nurture) part of the human. The more that is known about the function and interrelationship between genes, the eventual physical human they result in, and how they influence social interaction, the more connections that can be made from each single piece of information. This is true about the individual from whom that information is drawn and about the genetic groups to which they belong. Further, as analysis technologies and understanding of genetics develop over time, interpretation of genetic data will become more advanced. Accordingly, the relevance of any piece, or collection, of genetic data will depend on who is processing it, in which context and when. This will define what information they are capable of extracting and how this information
can be interpreted. Information that is harmless, and even unrecognisable as genetically significant in one processing context, becomes highly sensitive in another. (Campbell, Reece, & Mitchell 1999, pp.224-388)

Genetic data could have relevance in almost all spheres of daily life and human activity. Having said this, there are certain areas of endeavour for which genetic data may be seen to be particularly relevant. Those that have hitherto received the most attention have been law enforcement, healthcare, medical and scientific research, employment and insurance. (Article 29 Data Protection Working Party 2004, pp.7-12)

2. Directive 95/46: Different Forms of Data, Different Levels of Protection

The current piece of legislation giving overarching practical expression to data protection is Directive 95/46. (European Parliament & European Council 1995). On the one hand, the Directive lays down obligations to the data processor for the legitimate processing of data. On the other hand, the data subject is invested with a series of rights – including the right to be informed about the processing of his or her data and the right to inspect or correct data being processed. The data subject does not always have the right to stop processing. The Directive constructs a legal framework according to which individuals’ personal data can be processed provided that a certain set of rules and principles are followed. As put by de Hert, “[the data protection framework] relates to procedural justice and to the correct treatment of, and explanation to, registered citizens with the intention to increase their willingness to accept a system in which others (government agencies, companies, private citizens) have the right to process ‘their’ data and take decisions that have an impact on their self-determination where information is concerned”. (De Hert 2009, pg.17)

One can consider the right to data protection as a transparency right. ‘[T]ransparency [rights]…come into play after normative choices have been made, in order to channel the normatively accepted exercise of power through the use of safeguards and guarantees in terms of accountability and transparency’. This form of right is to be juxtaposed to opacity rights – such as privacy – which seek to provide prohibitive, substantive protection and which outright define whether an interference with the individual is acceptable. (Gutwirth et al. 2011, pg.8)

Within this framework, a number of categories of data are recognised. The framework provides tiered levels of protection, classifying data according to how much of an impact their processing could have on the fundamental rights of the individual to whom they pertain.

The first category of data recognised is ‘anonymous data’. Article 3(1) elaborates the scope of the Directive stating, ‘[t]his Directive shall apply to the processing of personal data’, clarified by Recital 27, which states ‘the principles of protection shall not apply to data rendered anonymous’(European Parliament & European Council 1995). Anonymous data can be any type of data as long as their connection with an identifiable individual has been severed. It is presumed that, as there is no connection between these data and an individual, there are no individual interests to be balanced against other interests in processing. This form of data is thus made exempt from the protection offered by the Directive.
The second category of data is ‘normal data’. This category includes all forms of data which are identifiable and which are not otherwise specified in the Directive. The processing of these data is subject to the default protection regime.

The third category of data is ‘sensitive data’. In Article 8 of the Directive, a set of specific rules are laid out for the processing of ‘special categories of data’. A stricter set of conditions and a higher standard of oversight are applied to these categories as opposed to ‘normal data’. This is on the presumption that the misuse of these types of data could have more serious consequences for the fundamental rights of the individual. (Article 29 Data Protection Working Party 2011, pg.4) Should data be classified under Article 8, their processing is generally prohibited, subject only certain limited exceptions laid out in Article 8(2) to Article 8(5).

3. Genetic Data and Directive 95/46

There are no specific mentions of genetic data in the text of Directive 95/46. However, there has never really been any doubt that genetic data fell within its scope. This was expressly recognised by the Article 29 Working Party (the group responsible for the ongoing interpretation of the data protection framework at European level). (Article 29 Data Protection Working Party 2004, pp.5-6) The fact that the Directive makes no mentions of genetic data means they are legally categorized are valued in the Directive in the same way as any other type of data (‘normal data’). This implies, in the eyes of the Directive, that the processing of genetic data has no different form of relationship to fundamental rights than the processing of any regular form of data.

The efficacy of a legislative instrument depends on how well it fits the substance and context of regulation. If there is a change in the substance under regulation, it is possible that the instrument diminishes in application and becomes less suitable to achieve its legislative ends.

In the case of the Data Protection Directive, the substance of regulation was ‘data’. According to the Article 29 Working Party, ‘genetic data show in themselves characteristics which make them singular [in relation to other forms of data]’. They summarize these unique characteristics as follows. 1: Genetic data is at once unique, in that it allows one specific person to be identified, but also has relevance to groups (for example blood relatives – including showing biological links between individuals) and even for succeeding generations. 2: Genetic data are potentially unknown to the bearer (and which the bearer may not want to know) and are outside the control of that bearer. 3: Genetic data can easily be obtained from raw material, although this data may be of dubious quality. 4: Owing to developments in research, the data may reveal more information in the future and may be used by an increasing number of entities. (Article 29 Data Protection Working Party 2004, pp.4-5) To this list, a fifth characteristic could be added; the information extracted from genetic data is highly dependent on the interpretative framework applied to it.

The recognition of the applicability of the data protection framework *rationae materiae* led to the need to apply the Directive to a substance for which it was not necessarily designed. Accordingly, the following questions relating to the suitability of the Di-
rective, as a system aimed at providing protection for fundamental rights, can be asked. Do genetic data have the same form of relationship with an individual’s fundamental rights as other forms of data? If not, does this different form of relationship undermine the suitability of the rules provided in the Directive to protect fundamental rights in the processing of genetic data?

In relation to these questions, criticisms of the Directive fall into two categories. First, that the protection offered by the Directive does not reflect the sensitivity of genetic data and is insubstantial. (European Commission 2012a, pg.67) Second, that there are a broader range of interests in genetic data, which the construction of the Directive does not recognise, and thus ignores.

4. Genetic Data is Recognized as ‘Sensitive’, but not by the Directive

In Article 8 of the Directive, a set of specific rules are laid out for the processing of ‘special categories of data’. A stricter set of conditions and a higher standard of oversight are applied to these categories as opposed to ‘normal data’. This is on the presumption that the misuse of these types of data could have more serious consequences for the fundamental rights of the individual. (Article 29 Data Protection Working Party 2011, pg.4) This is a closed group and genetic data do not feature under it.

However, a range of other legal sources recognize that certain forms of genetic data display characteristics which quantify them as potentially serious threats to the fundamental rights of the individuals they represent. The United Nations Educational Scientific and Cultural Organization (UNESCO) recognizes the specificity and sensitivity of genetic data to the impact of human rights. In the ‘Universal Declaration on the Human Genome and Human Rights’ and ‘The International Declaration on Human Genetic Data’. (UNESCO 1998; UNESCO 2003) The latter explicitly stating that ‘[t]he aims of this Declaration are: to ensure the respect of human dignity and protection of human rights and fundamental freedoms in the collection, processing, use and storage of human genetic data’. (UNESCO 2003, Article 1(a)) It its introduction, the Declaration recognizes ‘that the collection, processing, use and storage of human genetic data have potential risks for the exercise and observance of human rights and fundamental freedoms and respect for human dignity’ and considers ‘that human genetic data have a special status on account of their sensitive nature’. (UNESCO 2003) The Council of Europe has also specifically addressed questions of the use of genetic data due to its perceived sensitivity, including in Articles 11-14 of the ‘Convention of Human Rights and Biomedicine’ and in its Recommendation on ‘Genetic Testing and Screening for Health Care Purposes’. (Council of Europe 1992; Council of Europe 1997, Articles 11-14) At the national level, the recognition of the sensitivity of genetic data has led to its use in various sectors being highly regulated, and even in some cases – such as insurance – prohibited. (Nys et al. 2002, eg. pg.17) Indeed, certain national legislators have used the space left to national competence in the Directive, to specifically recognize genetic data as ‘sensitive’ in their national transpositions. (European Commission 2012a, pg.13) However, perhaps the most important legal forum to have recognized the sensitivity of genetic data is the European Court of Human Rights. This is significant due to the close
connection of the right to privacy as laid out in Article 8 of the European Convention of Human Rights and interpreted in contemporary society by the ECtHR and the function of data protection law under Directive 95/46. Article 8 ECHR can be seen as a ‘constitutional’ source which imposes upon states the positive obligation to ensure that laws ‘provide adequate protection against the unjustified disclosure of personal information’. The Directive, and its consequent national transpositions, can be seen as expressions of this positive obligation. In this respect, when a form of data is specifically recognized as sensitive by the ECtHR there is a strong argument that this sensitivity should also be reflected in data protection legislation.

The Court unambiguously recognizes the sensitivity of genetic data in the *Marper v. UK* case (European Court of Human Rights 2008b). In this case the European Court recognizes that cellular samples have a *particular* potential to pertain to the private sphere on account of the information they potentially contain and the use to which this information could be put. The Court sees this information as ‘intrinsically private’ and recognizes a strong individual and community interest in the tight control of its collection and processing. (European Court of Human Rights, 2008b, §104) The Court notes that it is impossible to evaluate the development of genetics technology and that it is impossible to ‘discount the possibility that in the future private-life interests bound up with genetic information may be affected in novel ways’. (European Court of Human Rights, 2008b, §71) The Court also notes that the content and quantity of information is significant. The information contained is seen as significant to the individual in its ability to reveal information of a highly personal nature (such as health details). The information is also seen as significant in its communal element, for example in its capacity to reveal family ties. Elsewhere in the judgment, relating to fingerprints, but with applicability to genetic data, the Court recognizes the privacy infringing quality of unique information. (European Court of Human Rights, 2008b, §§70-86)

The Court’s evaluation of the sensitivity of cellular samples revolves around the sensitivity of the *information* they contain, not the samples themselves. Accordingly this is a conclusion which can be expanded to include all forms of genetic data which are capable of revealing broad quantities of information. It is also a conclusion that can be proportionally expanded to genetic data which only display certain of the features listed as privacy sensitive.

Further observations of the potential fundamental rights significance of genetic data are to be found in the sources on which the Court draws. The Nuffield Council of bioethics, for example, points out the possibilities of ethnic identification, profiling and discrimination possible through genetic data. (Nuffield Council on Bioethics 2007, pg.56-57)

The fact that genetic data is not mentioned in the Directive means its intrinsic privacy is not recognized in the overarching legal instrument of data protection. This silence demonstrates a lack of alignment with the declaration of the potential threat genetic data

---

4 http://www.publications.parliament.uk/pa/jt200708/jtselect/jtrights/72/7205.htm
5 Other commentators have pointed out a number of other risks to fundamental rights, including to the violation of human dignity implied by a consideration of aspects of the biology of an individual, as opposed to engaging with that individual themselves as a person. (French National Consultative Ethics Committee for Health and Life Sciences 2007, pp.3-13)
could pose to fundamental rights elaborated in the various related legislative endeavors and in particular by the ECtHR in the Marper case. This has not meant that the processing of genetic data has occurred in precisely the same way as ‘normal data’ and that individual’ rights have been frivolously infringed. National legislations and supervisory authorities have taken up the slack and instituted their own control mechanisms and proportionality checks. This ‘bottom up’ approach, however, has led to another form of problem – legal fragmentation in the rules applicable to genetic data.

Indeed, given the description of the sensitivity of genetic data in Marper, one could question whether even the more stringent set of rules laid out in Article 8 of the Directive would suffice, as they still allow many instances in which processing is possible. The Court describes at length the numerous characteristics of genetic data capable of infringing fundamental rights and accordingly describes the information potentially contained as ‘intrinsically private’. This lengthy description of the privacy qualities of genetic data appears more acute than the Court’s description of the privacy qualities of other forms of data which are classified as ‘sensitive’ under Article 8 of the Directive. For example, one can compare the language used in Marper, with that used in I v. Finland, where the HIV status of an individual was only referred to as ‘highly intimate and sensitive’ and even then only in the sense of a revelation of health information.(European Court of Human Rights 2008a, §48)6 If we expand this reasoning to a general conclusion that genetic data is regarded by the Court as even more ‘sensitive’ than the other forms of data falling under Article 8, a very significant question comes forward. Is a framework of procedural justice appropriate, or are the individual interests in genetic data, or the threat its processing poses to fundamental rights, so great, that a more substantive construction would be required?

5. Individual Interests in Genetic Data and the Scope of Directive 95/46

The Directive constructs a framework aimed at providing procedural justice for data subjects. For procedural justice to be legitimate and effective, the framework needs to be capable of recognizing all the rights and interests which could be affected, and all the parties to whom these rights belong, in any given processing situation. Any legal framework which ignores legitimate interests cannot be said to offer an effective protection for fundamental rights. The Directive employs certain notions, such as ‘personal data’ and ‘data subject’, which restrict the scope of protection to certain individuals, and to certain forms of data. If these criteria are not met, the protection offered by the framework does not to apply. The assumption is that these notions must represent all relevant interests in data. This is a questionable assumption in relation to genetic data.

First, the framework operates on the presumption that interests in ‘personal data’ are carried by one single ‘data subject’. The reconciliation of interests in the processing of

---

6 The author makes this comparison very tentatively as a comparison was not intended to be made by the Court. However, there is enough specific reference to the sensitivity of genetic data in the Marper judgment and in other legislation to justify the general question.
data focuses around the relationship between this data subject and the controller. It is toward this subject that the controller directs their obligations. It is this subject who is seen to hold all rights in their ‘personal data’. Genetic data challenges the presumption of the single data subject by simultaneously containing information referring to other individuals. (Article 29 Data Protection Working Party 2004, pp.4-5) If genetic data, as such, are regarded as intrinsically private, and certain of the information which makes this data sensitive (the example of health information is given) can also be extracted about other individuals to whom the data relate (such as family members), it seems logical to assume that these other parties could also be said to have fundamental rights interests in this data. Data protection legislation ought therefore to recognize these other individuals’ rights and interests in some form.

The possibility may theoretically exist to recognize multiple data subjects under the Directive – as was even explicitly recognised in a previous opinion by the Article 29 Working Party, which even postulated the existence of a new legally relevant social class; the biological group. (Article 29 Data Protection Working Party 2011, pg.9) However, the reality of recognising multiple data subjects would be somewhat more complicated. This would imply the recognition of a much more complex set of relationships between genetic data and individuals (rather than an individual) and a much more complicated set of interest conflicts in the control of this data. The Directive provides no answers to the following questions. How would rights be distributed, would all data subjects have the same rights, or control, over data? If not, how could rights be allocated? To what lengths would the controller need to go to define and contact data subjects? How would the controller know to whom, or to how many, data subjects, to discharge obligations? Would all data subjects demand the same consideration? What would the correct course of action be if the controller faced infringing one data subject’s rights by fulfilling obligations in relation to another? (Taylor 2012, pp.116-119)

Owing to the silence in the recognition of multiple data subjects, and the lack of construction of a framework which considers their interests, these interests are in practice, ignored.

In relation to the use of genetic data in research on hereditary conditions, we can see this problem in practice. For research involving non-anonymized genetic material (and the data extracted from this material) to be legitimate, the consent of the donor (the data subject) must be secured. This consent is supposedly the expression of the donor’s considered autonomous wishes to allow an interference with their private information, and accordingly, serves to justify a lifting of restrictions on the processing of any collected data. In this case, the extracted genetic data will very probably reveal sensitive health information which also pertains to that individual’s relatives. There is no mechanism for the consultation of these relatives and the research can go ahead without their consent or even awareness. At this point, the rights of the data subject’s relatives to informational self-determination (normally aimed to be secured through the legal tool of consent), autonomy and an awareness of who, when and why their information is being processed, have been overridden. The problem becomes even more acute should the individual grant permission for results relating to the research to be published – thereby releasing sensitive information relating to a number of people, into the public sphere. At this point, the relatives of the individual have had sensitive information about them essentially released into the public domain, entailing a _prima facie_ breach of their privacy.
They may also now be subject to a number of consequent privacy risks – for example, discrimination against them on the basis of their genetic traits.

Second, the framework relies on the presumption that interests in data arise through the existence of a certain connection between data, and an identifiable data subject. Accordingly, the protection offered by the Directive does not apply to anonymous data. This is explicitly recognized in Recital 26 ‘the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable’. (European Parliament & European Council 1995, Recital 26) Once data are decreed to be anonymous once, they can then be used across processing contexts, by multiple controllers and are not subject to the obligations laid out in the framework. The initial data subject looses all control and oversight of who, how, why or when their data is being distributed or processed. There is no more transparency and no more control.

The lack of protection extended to anonymous data rests on three presumptions. 1: That it is possible to sever the connection between data and individual. 2: Once a connection has been severed, the anonymity of data will be its enduring status. 3. That there are no interests worthy of protection in anonymized data.

Genetic data challenge the assumption that a link can be cut between data and an individual. Anonymity as used in the Directive has focused on the severance of the connection between data and a civil identity. Genetic data remain a biological representation of one single individual even if the link to the civil identity has been cut. (Schmidt & Callier 2012, pp.305-309) The data thus remain uniquely personal, regardless of what is done to them. This makes it difficult, if not impossible, to suggest that genetic information is anonymous in the same way as, for example, transport information.

Genetic data also challenge the assumption that anonymity can ever be an enduring status. First, genetic data remain at a personal level; second, as explained above, the information contained in genetic data is dependent on the interpretative framework applied to them. New information can be revealed from an original genetic data set as technology develops allowing new analysis. New information can also be extracted as the processing context changes and a new interpretation is applied, allowing a different set of connections and conclusions to be drawn from the data. Accordingly, genetic data can only realistically be considered anonymous in the circumstances in which their anonymity was originally declared. In relation to the ability to draw new connections and inferences from samples/data previously considered unidentifiable, we can take an example from the world of forensics. Previously, a certain size of sample was generally needed to generate an individual’s unique profile (samples of less than 1mm diameter could be difficult). Due to recent technological advances, techniques now exist allowing the generation of a unique profile from a much smaller quantity of material. Accordingly, samples which could once have been considered to be anonymous are now, through technological advance, substances which allow an individual’s unique identification (Nuffield Council on Bioethics 2007, pp18-19).

As genetic data undermine these assumptions, the use of the concept of anonymity as an applicable concept to define the scope of data protection law masks interests the original data subject may retain in data incorrectly labeled anonymous.

Finally, the possibilities offered by genetic data challenge the presumption that there are no interests present in the use of anonymized data. Data is defined as anonymous when
the unique characteristics linking it to one ‘identifiable’ individual are removed (very few jurisdictions recognize ‘identifiable’ as data referring to more than one person). Data can thus apply to a limited group of people, perhaps very limited, without engaging the protection offered by the data protection framework in relation to any single one of those people, or those people collectively as a group. First, this means that the framework cannot recognize any interests involved at the level of the group – which is a particular weakness in relation to genetic data, given the historic use of genetics to found prejudice. Groups can be said to be subject to two forms of potential harm: internal and external. Internal harm can occur as genetic research (and often now genomic research – the study of population genetics) may reveal facts which dispute a group’s inherited cultural history. For example if research does not find a genetic basis to the claim of an ethnic group, or demonstrates that certain members of that group do not have the same genetic characteristics. External harms come from judgment being passed upon that group as a result of genetic findings supposedly unique to that group. For example, members of the Ashkenazi Jewish community expressed concern that research relating to their above average genetic predisposition toward certain cancers would found opinions toward the community as a whole. (Wang 2011, pp 88-98)

Second, genetic analysis of a group has implications for each member of that group, on an individual level. Accordingly, processing of anonymous data can still give rise to personal effect, and accordingly engages personal interests.

6. Genetic Data in the Proposed Data Protection Regulation

In January 2012, the Commission released the ‘Proposal for a Regulation…on the protection of individuals with regard to the processing of personal data and on the free movement of such data’. (European Commission 2012b) This was the culmination of a data protection reform process in the form of a draft Regulation proposed as a replacement to Directive 95/46. The prevailing mood of the Regulation is not that of change, but that of continuity. The overall goals of the proposed Regulation remain unchanged from those of Directive 95/46. The proposed Regulation still seeks to protect the rights of the data subject and guarantee the free flow of data between Member States. In general construction, the Regulation still aims at the provision of procedural justice through a framework in which different interests can be balanced. In content, the Regulation retains all the concepts and principles and many of the definitions, which defined Directive 95/46. (European Data Protection Supervisor 2012, pp.2-4) Despite the prevailing emphasis on continuity, the Commission did decide to innovate in relation to the regulation of genetic data.

According to the Article 29 Working Party, it is feasible that anonymous data could (re)become ‘personal data’ if the result of the processing impacts on that individual. (Article 29 Data Protection Working Party 2007, pg.10) However, it is not certain if the result needs to be specifically targeted at this individual. Further, the reassociation process would only be valid for the processing context in which it had occurred. This would not return control to the individual of the same anonymized data distributed elsewhere.
Article 2 defines the scope of the proposed Regulation. For genetic data to fall under the scope of the Regulation, they must still qualify as ‘personal data’ under Article 2(1). This is confirmed when read in conjunction with Recital 23, which states that ‘[t]he principles of data protection should not apply to data rendered anonymous in such a way that the data subject is no longer identifiable’.

In Article 4, the Commission, in defining ‘data subject’ state: ‘data subject’ means an identified natural person or a natural person who can be identified, directly or indirectly, by means reasonably likely to be used by the controller or by any other natural or legal person, in particular by reference to one or more factors specific to…the genetic…identity of that person. This reference sets the scene for the specific regulation of genetic data.

The decision to name, and to regulate, genetic data specifically as a part of broader reform, comes with the requirement to define genetic data. Accordingly, in Article 4 the Commission provide definition: ‘[g]enetic data’ means all data, of whatever type, concerning the characteristics of an individual which are inherited or acquired during early prenatal development’.

If we return to the explanation of genetic data given in part 1, we can see that the phrase ‘all data, of whatever type, concerning the characteristics of an individual which are inherited…’ could be extended to a potentially seismic quantity of data. This would certainly, always, include any genetic test results looking at the specific genetic architecture of the individual. This could also be seen to include all information about an individual which reflects the genetic architecture physically or socially, or indeed any consequent social interpretation of these. Data such as ‘age’, ‘height’ or ‘sex’, could very easily be seen to fall under this definition, as could all photographs.

There is also no reference to the context in which data must be processed to fall under the definition of ‘genetic’ offered. As we see in part 1, many forms of data may contain a genetic element, yet the significance of this genetic element may vary depending on the intention of processing – we may not think of the use of eye colour, as an identifying feature, as implying the processing of genetic data, but we would certainly consider the use of eye colour, processed to screen for hereditary disease, as the processing of genetic data. The lack of reference to processing context suggests that there does not even need to be an intention to interpret data genetically to still qualify those data under this category – genetic data are de facto always genetic data. In general, the current definition seems overly broad and it would seem likely that it will be amended.

A further point that should be made is that, in defining genetic data as a category, the Commission must mean that genetic data is to be differentiated from health data, and indeed from any other specifically named form of data (biometric data, for example). The precise dividing line will need to be clarified. Under the Directive, all sensitive data were treated the same. Under the Regulation, different forms of sensitive data may engage different grounds for legitimation under Article 9 (for example Article 9(2)(h) for health data) and engage potentially different regimes – see the engagement of Article 81 for health data and Article 83 for sensitive data in the research context. A conflict is

---

8 An interesting question would thus be: Who defines the interpretation relevant to define data as ‘genetic’?
thus feasible between which regime is engaged. The logical assumption would be that
the regime offering most protection would need to be chosen. However, this is not clari-
fied in the Regulation, nor is it always certain how such a decision could be made. 9

Should data fit within the definition of genetic data in Article 4, this has the effect of
categorizing them amongst the ‘sensitive’ categories of data elaborated in Article 9. The
processing of data classified as sensitive is, in principle, prohibited – Article 9(1) ex-
plicitly stating ‘The processing of personal data, revealing race or ethnic origin, political
opinions, religion or beliefs, trade-union membership, and the processing of genetic
data…shall be prohibited’.

Article 9(1) employs a very peculiar construction that deserves further consideration.
The Article appears to split the forms of sensitive data into two categories; 1. ‘the pro-
cessing of personal data, revealing race or ethnic origin, political opinions, religion or
beliefs, trade-union membership’ and 2. ‘The processing of genetic data…’ A literal
reading of this wording would suggest that the ‘genetic data’ do not need to fulfill the
criteria of also being ‘personal’. This would imply a blanket ban on the processing of all
genetic data, including anonymous genetic data. There is no clarification of this word-
ing in the explanatory note accompanying the release of the Regulation, nor in any of
the available preparatory documentation. However, to include even non-personal genet-
ic data would be a novel step. This lacks precedent elsewhere in the Directive or Regu-
lation – both of which have relied on the concept of ‘personal data’ as foundational el-
ements – and it seems strange that the Commission would take such a step without mak-
ing further comment of explanation. As the correct interpretation remains somewhat
unclear, both interpretations remain plausible (1. genetic data do not need to be ‘perso-
nal’ to qualify under Article 9, 2. genetic data must also be personal). The significance of
this uncertainty will be returned to below (in part 8).

This general prohibition is qualified by a series of exceptions laid out in Article 9(2).
Should a processor wish to process genetic data, the processing must be legitimated by
one of these exceptions. The exceptions according to which genetic data will most
commonly be processed (under current conditions) are specified in 9(2)(a) – consent,
and 9(2)(i) – scientific, historical or statistical research purposes. 10

For a consent to be legitimate, 9 (2)(a) states that it shall be ‘subject to the conditions
laid down in Articles 7…except where Union law or Member State law provide that the
prohibition referred to in paragraph 1 may not be lifted by the data subject’. Consent is
defined in Article 4 of the Regulation as ‘any freely given specific, informed and explic-
it indication of his or her wishes by which the data subject, either by a statement or by a
clear affirmative action, signifies agreement to personal data relating to them being pro-
cessed’. The specifics of this explicit consent are clarified in a series of other recitals,
most relevantly Recital 25. Article 7 elaborates further formal features of consent, per-
haps the most interesting of which is in 7(4), which states that consent shall not ‘provide
a legal basis for processing…where there is a significant imbalance between the posi-
tion of the data subject and controller’. This is clarified in Recital 34 as ‘especially the

---

9 This is a theoretical proposition at present. As interpretation of the Regulation begins, the reality of the
problem will reveal itself.
case where the data subject is in a situation of dependence from the controller, amongst others, where personal data are processed by the employer’.

The exception in Article 9 2(i) is clarified in Article 83 – under the general heading of ‘Provisions Relating to Specific Data Processing Situations’. Article 83 states that genetic data could only be processed for these purposes if, 1: the purposes followed cannot be achieved by processing anonymised data, or 2: ‘data enabling the attribution of information to an identified or identifiable data subject is kept separately from the other information’ as far as is possible’. The publication of research relying on personal genetic data is permitted if; 1: consent has been granted by the data subject, 2: the publication of research findings relies on this and the publication would not interfere with the fundamental rights of the data subject; 3: the genetic data have already been made manifestly public by the data subject. A number of recitals clarify this article, defining rather broadly the limits of the term ‘scientific research’, for example Recital 126. Still other recitals, such as 40, 42, 50, 53 and 88 highlight provisions still applicable to the exception, such as that of purpose specification.

A series of other exceptions may become functional in certain situations, for example, 9(2)(g) – in the carrying out of a task in the public interests or on the basis of Member State or EU law. Other exceptions may become important (in quantity of use or in controversy) depending on how the processing of genetic data develops, for example Article 9(2)(e) – data which have been manifestly made public by the data subject. One thinks, for example, of a scenario in which people begin to publish their genetic data, engaging this exception when the data are used.

Article 9(3) states that the Commission shall be empowered to adopt delegated acts to clarify the criteria under which, and safeguards according to which, sensitive data can be processed.

Having a legitimate ground under 9(2) on which to process genetic data does not release the controller from fulfilling other obligations for processing laid down in the Regulation. These are predominantly laid out in Article 5. Genetic data must be processed fairly, lawfully (defined in Article 9) and transparently in relation to the data subject (defined in Articles 10-13). They must be collected for specific, explicit and legitimate purposes, and not further processed in ways incompatible with those purposes. They must be relevant to the purpose for which they were collected and limited and necessary for that purpose. They must be kept accurate and up to date and must be kept in a form which only allows identification of data subjects for as long as absolutely necessary. The controller is also obligated to provide the data subject with information related to the processing of their data (laid out in Article 14).

Nor does having a legitimate ground for processing necessarily divest rights from the data subject. While legitimation under any other ground than consent or manifest publication would legitimately override the data subject’s right to informational self-determination, there are a series of other rights which may still remain valid. These are predominantly laid out in Chapter III, ‘Rights of the Data Subject’. Article 15 provides that the data subject has a right to information as to their genetic data being processed and as to the certain details of that processing. Articles 16 and 17 provide rights to rectification and erasure of incorrect, or improperly processed or retained data. Article 20 gives every natural person the right not to be subject to measures based solely on automated processing.
Should the data processor wish to process genetic data they may also be subject to the requirements laid down in Article 33. Under this Article, the data controller is obliged to carry out a Data Protection Impact Assessment (DPIA) to evaluate the processing operation for potential impacts on fundamental rights and freedoms. A DPIA will certainly need to be carried out if the controller wishes to have a large scale filing system of genetic data as this is a type of processing specifically pointed out to ‘present specific risks’ under Article 33(2)(d). However, many more forms of processing of genetic data could require a DPIA. Article 33(1) states; ‘where processing operations present specific risks to the rights and freedoms of data subjects by virtue of their nature, their scope or their purposes’ a DPIA shall be carried out. In classifying all genetic data as sensitive under Article 9 – and thus recognizing them as data whose misuse could have severe consequences on the individual’s fundamental rights – the Commission have also imply that much processing involving genetic data fall under the criteria laid out in Article 33(1).

In the case where a DPIA has revealed specific risks related to the processing of genetic data falling under 33(1) or 33(2)(d), processing will be subject to prior consultation with the DPA under Article 34(2)(a). The same requirement for prior consultation will arise if the processing operation has been listed by the DPA as ‘likely to present specific risks to the rights and freedoms of data subjects’ under Article 34(2)(b) and 34(4). Given that genetic data have been defined as ‘sensitive’, the attention DPAs have paid to genetic data up until now and the uncertainty in the consequences of its processing, it would seem likely that many processing operations involving genetic data will also be on the lists drawn up by national DPAs.

Finally, should processing of genetic data occur in violation of the rules of the Regulation, a competent DPA would have the power to impose sanctions as laid out in Article 79 (1). Article 79(2) states that these sanctions must be ‘effective, proportionate and dissuasive’ whilst Article 79(6) states, that the ‘fine [could be] up to 1 000 000 EUR or, in case of an enterprise up to 2 % of its annual worldwide turnover, to anyone who, intentionally or negligently...[Article 79(6)(b)] processes special categories of data in violation of Articles 9’ or (Article 79(6)(i)) ‘does not carry out a data protection impact assessment pursuant or processes personal data without prior authorization or prior consultation of the supervisory authority pursuant to Articles 33 and 34’. (European Commission 2012b)

7. The Regulation Recognizes the Sensitivity of Genetic Data

It seems relatively clear that, in relation to the protection offered by the Directive (if a data subject can be identified) data subjects’ rights in relation to the processing of their genetic data will have been strengthened.11 On a broad level, the proposed Regulation generally strengthens data subjects’ rights over their personal data and demands greater transparency and accountability from pro-

---

11 A different question could also be asked: does the protection offered by the Regulation to the current protection standards in the Member States? The Regulation does not necessarily aim at harmonizing at the highest level of protection and it is possible that the protection standard offered to genetic data in some Member States is in fact higher than that proposed under the Regulation.
cessors. Whilst these features are not specific to genetic data, these rights and obligations will apply to the processing of genetic data (unless excluded) and accordingly are an important part of the broader system defining interests in data and under which genetic data processing will take place.

However, it is in the addition of genetic data to the list of ‘sensitive data’ that the Regulation takes the biggest step in protecting fundamental rights threatened by the processing of genetic data. It is here that the Regulation directly addresses the silence, and remedies the lack of recognition in the Directive.

The purpose of the ‘sensitive’ category in the Regulation follows the same logic as that of Article 8 in the Directive, stemming from ‘the presumption that misuse of these data could have more severe consequences on the individual’s fundamental rights, such as the right to privacy and non-discrimination, than misuse of other, “normal” personal data’ (Article 29 Data Protection Working Party 2011, pg.4). Accordingly, when data are identified as ‘sensitive’ in the Regulation, they engage a more stringent set of rules for processing than ‘normal data’ (De Hert & Papakonstantinou 2012, pp.132-133).

Indeed, under Article 9 of the Regulation, the standard of protection may even be superior to that offered under Article 8 of the Directive. Article 9 offers protection for fundamental rights in three ways. First, the formulation of Article 9 is that of general prohibition on processing. This can be opposed to the general ‘permission with conditions’ formulation for normal data. ‘Sensitive’ data can only be processed provided this processing falls under a specific exception laid out in 9(2). This means that a number of forms, reasons and contexts of processing possible with ‘normal data’, become illegitimate with sensitive data. This construction is much closer to offering substantive protection than that applying to ‘normal data’. With specific focus on the classification of genetic data as sensitive, one commentator makes the relevant observation, ‘[i]n the case of the extension of the category of “sensitive data”, a possible positive outcome could consist of the de-trivialization of the storing and processing of DNA data’ (Bellanova et al. 2012, pg.116).

Second, if a legitimate reason for processing is found under 9(2), the processing can only proceed following the discharge of greater burdens and more stringent oversight; first, through the DPIA mandated by Article 33 and second, through the mandated advance oversight and checking of the processing by the DPA in Article 34. These mechanisms are designed to allow further accountability, transparency and oversight of each case of processing so that the fundamental rights of the data subject can be considered in more detail on a case by case basis. Any instance of processing which would disproportionately infringe these rights can then be detected and prevented, or amended as necessary.

Third, the Regulation contains an inbuilt dissuasive/punitive element to the negligent or wilfully illegitimate processing of genetic data. It is perhaps here where the Regulation differs most from the regime under Article 8 of the Directive. This manifests predominantly in the fines system under Article 79. 1 million Euros, or 2% of annual worldwide turnover is not an insubstantial amount. The power given to DPAs to levy this fine will certainly act to motivate care in the discharge of obligations on the processing of genetic data. However, whether this Article remains unchanged until the end of the reform process is highly uncertain. The proposed fines system has already been the target of significant debate and lobbying.
What the Regulation still does not provide however, is **absolute** substantive prohibitions on processing. This was not the goal of the Directive and is not the goal of the Regulation. Both instruments work on the basis that data can be legitimately processed and aim at controlling and channelling this legitimate processing. As explained above, they both follow a transparency logic – not designed for prohibitions or preventions. However, given the strong language used by the ECtHR relating to the unique privacy and intimacy of genetic data, as well as the host of specific legislation it has engendered, an argument could be made that even the ‘alert’ and supervision system offered by the Regulation is too lenient. This argument would suggest that an alternative approach for the processing of genetic data should have been instituted. (European Economic and Social Committee 2012, pg.93) The alternative approach available to the European legislator would have been to establish substantive protection, and draw firm limits on the possibilities to infringe into the individual’s private sphere through the processing of genetic data – an opacity approach. The authors find this argument unfair, and applaud the Commission’s restraint. De Hert and Gutwirth observe that, whilst a combination of both forms of tool (opacity and transparency) is necessary to ensure optimal regulation of new technologies – and the balance of this choice must be constantly revisited – a number of criteria can be considered as to which tool should best be employed in each situation. They consider ‘opacity/privacy rules - prohibitionary rules – should guarantee these aspects of an individual’s life that embody the conditions for his/her autonomy (or self-determination, or freedom, or ‘personal sovereignty’)’ whilst ‘[i]n all cases where consent (still) plays an important role, it can be assumed that the guidelines for the opacity approach are not met’. (De Hert & Gutwirth 2003, pg.151,153) They clearly suggest the deployment of opacity tools to be a more heavy-handed measure and state, ‘[a]pproaching new phenomena with heavy prohibitions may lead to situations in which the prohibitions are not respected or technological development is blocked’ (De Hert & Gutwirth 2003, pg.93). It may well be that the sensitivity of genetic data eventually proves too dangerous for individual liberties to allow any form of transparency approach. However, for now, that point has not been reached and should the Commission have laid down substantive rules, the result may have been counterproductive.

We can be more specific in this argument. First, the Commission would have had to exceptionralise genetic data, giving it a form of protection beyond that of all other forms of data. They would thus have been forced to conduct a difficult rationalisation as to why genetic data were more privacy invasive than, for example health data, data relating to ethnicity or biometric data. Depending on context, many other forms of data can be far more privacy sensitive than many forms of genetic data. Knowledge of an individual’s HIV status, for example, is far more privacy sensitive than, for example, knowing an individual’s eye colour. Further, the Commission would have been forced to draw substantive lines of protection. How could a decision be made relating to which genetic data, or forms of processing, to prohibit? There are very few legal sources which could clarify how, or where a line could be drawn – which precise aspects of genetic data processing interfere with the conditions for individual autonomy, and how? The Marper case is perhaps the most famous case directly verifying the possible impact of genetic data on the right to priva-
However, despite listing numerous qualities which make genetic data sensitive, and confirming its unique sensitivity, the case only draws the vaguest of conclusions as to when genetic data actually infringe privacy or how their use should be balanced against other interests such as research or public health. (European Court of Human Rights 2008b) Accordingly, the Commission would have had to define the specifics of this substantive protection. They would have needed to define which forms of genetic data could be processed and which were prohibited and to state why. The linked nature of the levels of genetic data (architecture describes feature and vice versa, providing indication about third connected trait...) would make it hard to justify the prohibition of one level (for example direct DNA analysis) whilst leaving the processing of other levels (examination of the phenotype, for example) legitimate. Prohibition of the processing of all data with a genetic element would have been equally impossible. The range of data types necessary to prohibit would be simply too large. Equally, the declaration of any general prohibition at European level would have had effect across numerous processing contexts, each with a unique set of interests involved. Processing of certain forms of genetic data in the context of health research could not reasonably be subjected to the same form of prohibition as the processing of genetic data by insurance companies. Even less could the Commission assume that they were centrally in the best position to elaborate these boundaries for all sectors (McNally et al. 2004).

Finally, one can consider the fact that substantive protection is not ruled out by the Commission’s approach. In the drafting process, three policy options for the Regulation were considered. 1: A non-legislative approach aimed at the clarification and application of existing rules via soft law. 2: Certain non-legislative amendments considered in option 1, plus a modernization and harmonization of the current framework. 3: A broad legislative approach consisting of detailed, prescriptive rules even extend to regulating specific sectors – a ‘centralized approach’. The final proposed Regulation predominantly follows the logic of option 2. This means that the rules provided by the Regulation are intended to be fleshed out and clarified by sector specific rules (European Commission 2012a, pp.44-62). Substantive protection can still be elaborated at sector specific level and developed according to a closer inspection and proportionality review of the interests involved in each sector, or processing context.

8. Limits in the Scope of Application Mean Interests May Remain Invisible

Despite the generally positive effect on the level of protection of fundamental rights by the definition and categorization of genetic data, the second form of problem remains unaddressed: The inability of the data protection framework to recognize the broader range of interests data subjects may have in genetic data.

The Regulation retains the concept of ‘personal data’, as a gateway to accessing the procedural justice offered by data protection law and there is no explicit mention of possible multiple data subjects, nor of any mechanism aimed at reconciling the above mentioned unaddressed interest conflicts (between data subjects and between a controller and multiple data subjects).
The sole recognition of rights for non-data subjects comes in Article 20 (clarified in Recital 58) relating to ‘[m]easures based on profiling’. Article 20(1) states; ‘Every natural person shall have the right not to be subject to a measure which produces legal effects concerning this natural person or significantly affects this natural person’. This is at least seems a recognition of the potential for third party interests to be touched by data processing. However, the Article was not designed to deal with issues arising from multiple data subjects or from genetic data. In the explanatory note accompanying the proposal, the Commission also states that the protection is indeed aimed at ‘data subjects’ rather than third parties (European Commission 2012b, pg.9). They also make clear that the ‘profiling’ under consideration relates to the possibilities presented by ICT and large scale data collection and combination, as opposed to genetic profiling. (European Commission 2012b, pg.9) Whilst time may demonstrate the relevance of this Article to genetic data, it is currently of uncertain application.

Given the lack of explicit legal recognition, there is no reason to assume the function of the framework under the Regulation would be any different to that under the Directive. Whilst multiple data subjects are not in principle excluded from recognition, the continued focus on ‘personal data’ implies a continued focus on the single data subject. This means that, in practice, the interests of other data subjects continue to be ignored.

The Regulation also retains the concept of ‘anonymous’ data and continues the exclusion of anonymized data from its scope of protection. There are no clear indications that the assumptions around ‘anonymity’ in the Directive have been reconsidered (that the link between data and individual can be severed, that this severance can be perpetual and that there are no worthy interests in anonymous data). The Regulation employs a wide definition of the term ‘identifiable’, including, ‘direct, or indirect identification’ taking into account ‘all means likely reasonably to be used’ by anybody toward the end of identification. (European Commission 2012b, Recital 23) However, this is not really a departure from the approach used in the Directive. Within this definition, the controller will need to consider the means being made available by future technological progression. (Article 29 Data Protection Working Party 2007, pg.15) This broad definition is clearly an attempt to set the bar for anonymity high. However, the evaluation as to anonymity will still be made by one controller from that controller’s own perspective. Given that the future content and development of technologies relating to genetic data is uncertain, the evaluation of which means are ‘reasonably likely’ would involve impossible calculation. The only possible solution to this calculation would be to assume that the anonymization of genetic data was in principle impossible. This is clearly not the intention of the Regulation. The casting of a broad net for identifiability is positive, but it leaves the concept of anonymity in relation to genetic data intact. Accordingly, it leaves the fundamental problems related to the anonymity of genetic data untouched.

---

12 The Commission refers to ‘data subject’ in the explanatory note, but to ‘natural person’ in the text of the Regulation. These are not the same thing. It could be that this is a mistake. It could also be that the Commission mean that a ‘natural person’ qualifying for protection under Article 20 becomes a ‘data subject’ due to the effect those data are having on that person.
The Regulation also recognises a form of context specific anonymity and the possibility for data to reveal more information depending on the context of their interpretation (pseudonymity). Article 83(2)(b) states, ‘data enabling the attribution of information to an identified or identifiable data subject is kept separately from the other information as long as these purposes can be fulfilled in this manner’. (European Commission 2012b) This would allow for broader use possibilities within data, whilst avoiding the associated problems related to anonymous data shifting across contexts and purposes. However, there are limits to the significance of this recognition. Article 83 only applies to the use processing of genetic data for historical, scientific or statistical purposes. Further, the recognition of the possibility of pseudonymity seems not to be aimed at rectifying potential problems in the concept of anonymity, but at providing an option for researchers for whom the use of anonymous data (in the sense of data protection law) is unfeasible. In fact, considering that Article 83(2)(a) promotes the use of anonymous data before pseudonymous data, it would seem that the use of pseudonymous data is seen as a secondary option to the use of anonymous data.

However, it is in the broader possible construction of Article 9(1) – that genetic data do not need to be personal to qualify for protection (see part 6) – where the Regulation offers the most hope in addressing the problems raised by the concept of anonymity. If this interpretation is correct, the Regulation may have engineered a situation in which the concept ceases to be as problematic. In this construction, the possibility to anonymize genetic data is not challenged, but interests in the processing of ‘anonymized’ genetic data are recognized and therefore the processing of genetic data qualifies for the protection offered by Article 9. This would not ensure the same level of protection as given to ‘personal’ genetic data, as the transparency of processing in relation to, and rights of the individual over, ‘personal’ genetic data would not be engaged. However it would mean that the Regulation’s other safeguards would be engaged, including external oversight (Articles 33 to 34), the obligation to process the data lawfully (Article 6) and the possibility for negligent/wilfully illegitimate processing to be severely punished (Article 79).

Accordingly, it is not at question that genetic data can still be ‘anonymized’ according to the Regulation. How problematic this concept remains will depend on the interpretation of Article 9 applied. If the broader interpretation is applied, and genetic data do not need to be ‘personal’, the processing of genetic data will qualify for a certain level of protection (how much, and whether this is effective are other interesting questions). If a narrower interpretation is applied, the problems related to the use of the concept of anonymous data in relation to genetic data will remain unaddressed.

**Conclusion**

Genetics is a fascinating and poignant topic. The possibilities offered by genetic analysis and insight present totally new challenges to our perception of the human, and accordingly to our perception of society. How law faces these challenges today may define how European society looks tomorrow.

Despite being one of the key areas of law dealing with the regulation of genetic data, the current EU data protection framework was not designed with genetic data in mind and
makes no specific reference to it. This led to questions as to whether the framework was able to achieve its core goal of providing protection for fundamental rights in relation to the processing of genetic data. First, it was seen not to recognise the inherent ‘sensitivity’ of genetic data. Second, its construction shut out recognition of a number of legitimate interests. It did not reckon with the possibility of multiple data subjects, or the implications of data which could reveal varying content and quantities of information depending on processing context.

The approach of the Regulation addresses the first issue well. The structure of the Regulation generally strengthens individual rights and in regulating genetic data under Article 9, it addresses the silence of the Directive and directly recognises the ‘sensitivity’ of genetic data. Processing of genetic data can now take place only under explicitly limited conditions, must be considered for its impact on fundamental rights and is subject to increased oversight and transparency requirements. Negligent or wilfully illegitimate processing can be punished with harsh fines levied by independent DPAs. Whilst the Regulation does not provide substantive protection against the processing of genetic data, this is, at the current juncture, welcome. Substantive protection in a general instrument at European level would be a sweeping move which would be very difficult to justify and would have the effect of obstructing numerous positive uses of genetic data. Furthermore, the current formulation of general rules following a transparency approach deliberately leaves room for specific prohibitions to be enacted on a sectoral level, where the consequences of such a prohibitionary/opacity approach can be better quantified and justified.

Unfortunately, it is in relation to the second issue; recognition of the range of interests in genetic data, where the reform proposals are more shaky. By retaining the concepts, and many of the definitions of the Directive, the Regulation left little space for the recognition of the more complex set of interests in genetic data. The concepts of ‘data subject’ and ‘personal data’ continue to recognise only the single data subject, in whom rights are invested and towards whom obligations must be addressed. These continue to act as a bar to the recognition of members of a data subject’s family, or biological group – to whom genetic data may also relate. The Regulation also retains a concept of ‘anonymity’ that seems basically unchanged from the Directive. Generally speaking, this implies that the Regulation retains certain further assumptions (that the link between data and individual can be severed, that this severance can be perpetual and that there are no worthy interests in anonymous data). These assumptions are not valid in relation to genetic data. First, as the information content of genetic data is dependent on who interprets it and when, the assumption of anonymity can only be relevant for the context in which that anonymity was declared. Second, even when data is declared anonymous, its use may still have an impact on the biological groups to which the individual belongs, and by proxy on the individual themselves. How problematic this concept remains will depend on the, currently ambiguous, application of Article 9. If a broader interpretation is applied, the Regulation will provide an exception to the default application of anonymity which would mean genetic data do not need to be ‘personal’ to qualify for protection. In this case the Regulation challenges the idea that there are no interests in the processing of ‘anonymized’ genetic data. If a narrower interpretation is applied, the problems related to the use of the concept of anonymous data in relation to genetic data will remain unaddressed.
References

European Court of Human Rights, I v. Finland, Application no. 20511/03. 2008a.
European Court of Human Rights, S. and Marper v United Kingdom, App no. 30562/04 and 30566/04. Marper. 4-12-2008b.
European Data Protection Supervisor 2012, Opinion of the European Data Protection Supervisor on the data protection reform package.
European Economic and Social Committee 2012, Opinion of the European Economic and Social Committee on the General Data Protection Regulation.

Gutwirth, S., Gellert, R., Bellanova, R., Friedewald, M., Schütz, P., Wright, D., Mordini, E., & Venier, S. 2011, Legal, social, economic and ethical conceptualisations of privacy and data protection D1.


Wang, Y. 2011, Group protection in human population genetic research in developing countries: the People's Republic of China as an example, Glasgow University.

Dara Hallinan & Michael Friedewald (dara.hallinan@isi.fraunhofer.de & michael.friedewald@isi.fraunhofer.de) Fraunhofer Institute for Systems and Innovation Research; Breslauer Straße 48, 76139, Karlsruhe, Germany

Professor Paul De Hert (Paul.De.Hert@vub.ac.be) Vrije Universiteit Brussel; Pleinlaan 2, 1050, Brussels, Belgium

13 Corresponding Author: Tel. No. +49 721 6809-369