11. JURISDICTIONAL CHALLENGES RELATED TO DNA DATA PROCESSING IN TRANSNATIONAL CLOUDS

Heidi Beate Bentzen* and Dan Jerker B. Svantesson**

1. INTRODUCTION

Genetic research has the potential to change how we diagnose, prevent and treat medical conditions, by making the diagnosis more precise and the prevention and treatment more personalised. However, such research cannot be carried out without the collection, use and disclosure of sensitive data – our DNA. Furthermore, to be effective, such research currently depends on DNA data being shared across borders and processed in cloud computing arrangements. Thus, genetic research is global, but it is not regulated similarly across the world.

In this chapter, we examine the jurisdictional issues that arise in both private, and public, international law, where DNA data is stored or processed in transnational cloud computing arrangements. Further, the broad contours of a potential approach to dealing with those issues will be canvassed. First, to set the scene for that discussion, we will commence with a brief discussion of what types of data we are dealing with here, what they are used for and the role cloud computing plays in the processing.

* Centre for Medical Ethics, Faculty of Medicine; Norwegian Research Centre for Computers and Law, Faculty of Law, both at the University of Oslo. Bentzen collaborates with the Norwegian Cancer Genomics Consortium. E-mail: h.b.bentzen@medisin.uio.no. Her research is financed by the Research Council of Norway through the project Legal Regulation of Information Processing relating to Personalised Cancer Medicine (BIOTEK2021/238999).

** Co-Director, Centre for Commercial Law, Faculty of Law, Bond University (Australia); Visiting Professor, Faculty of Law, Masaryk University (Czech Republic); Researcher, Swedish Law & Informatics Research Institute, Stockholm University (Sweden). E-mail: Dan_Svantesson@bond.edu.au. Professor Svantesson is the recipient of an Australian Research Council Future Fellowship (project number FT120100583). The views expressed herein are those of the author and are not necessarily those of the Australian Research Council.

The authors wish to thank the two anonymous reviewers for their useful feedback on this chapter. In addition, the authors would like to thank Professor Dag Wiese Schartum and PhD candidate Isabelle Budin-Ljøsne for their helpful comments.

Our aim is modest. We do not aim to be exhaustive on any one topic. Rather, by providing a brief introduction to both DNA data processing and to the legal issues, we aim to make the presentation accessible to a diverse audience, hopefully making this chapter a suitable starting point for anyone considering researching in detail the cross-section of transnational DNA databases and international law.

2. DNA IN THE CLOUDS – THE BASICS

2.1. HOW AND WHY DNA DATA IS USED

The human genome refers to an individual human being's complete set of DNA. In the current European Data Protection Directive 95/46/EC, genetic data is commonly considered health data, which is characterised as sensitive personal data. In the upcoming European Union General Data Protection Regulation (EU) 2016/679, genetic data is explicitly regulated as a special category of data alongside, *inter alia*, health data. We will return to the legal regulation in section 4 below.

Genomic data is a unique identifier. Furthermore, an individual can be identified even by very little DNA data. Anonymous processing is therefore rarely an option, so the processing must be in compliance with the applicable personal data legislation. Some also take the view that the characteristics of genomic data sets it apart from other kinds of sensitive personal data, and that the processing requires even more consideration than for other types of sensitive personal data.[1]

DNA testing can be used for various purposes, ranging from personalised medicine and parentage identification, to ancestry research and sport talent identification.[2] There has also been a proliferation of direct-to-consumer (DTC) genetic tests, which has created concerns, not least due to the fact that the laboratories carrying out the DTC tests often are located in a different country to where the consumer is located. And immediately, we see the type of (cross-border) data privacy concerns that arise in the field that we focus on in this chapter. For example, Australia's National Health and Medical Research Council has pointed out that: 'Some DTC companies also sell information about you and your genetic results to pharmaceutical and other companies. It is important

1 According to the UNESCO Declaration on Human Genetic Data, Art. 4 genetic data have a 'special status'.

2 NHMRC, 'Use of genetic infromation in sport', 2013, <https://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/g003_genetic_sequencing_in_sport_150622.pdf> accessed 08.07.2016.

to understand that DTC genetic testing companies may ask if your sample and results can be used for other purposes, such as research.'[3]

At any rate, a particularly important area of DNA use is linked to so-called 'personalised medicine'. Personalised medicine is the tailoring of prevention, diagnosis and treatment to each individual's DNA. The aim is to be able to identify the most effective treatment, decrease the time it takes for patients to be given that effective treatment, and minimise side-effects.

Of the various types of DNA test, the one that is usually considered most relevant to personalised medicine is genome sequencing. Genome sequencing maps an individual's entire DNA, including all the genes and all the non-coding regions. That means that all of the about 3.2 billion base pairs that constitute the genome are mapped. The genes themselves only make up a small portion of the DNA, about 1.22 per cent.[4] The non-coding regions are by far the biggest part of the DNA. These are interesting because they play a part in gene regulation. Other available DNA tests are exome sequencing, that maps all the genes but not the non-coding regions, gene tests that only map select genes, and DNA tests usually used for identification purposes that only map small parts of the non-coding regions.[5]

The opportunity to tailor medical care to the individual in the manner done in personalised medicine is new. It has been made possible by advances in medical research and technology. In 2000, a rough draft sequence of the human genome was finished, and the achievement was announced jointly by US President Bill Clinton and UK Prime Minister Tony Blair, before it was published in 2001.[6] In 2004, the first complete human genome sequence was published.[7] These are considered to be among the main medical research achievements in history. Simultaneously, technological advances have rapidly made genome sequencing affordable. It cost about $3 billion to sequence the first human genome. In October 2015, the cost had fallen to $1,245.[8]

[3] NHMRC, 'Understanding Direct-to-Consumer (DTC) Genetic DNA Testing: An information resource for consumers', 2014, <https://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/g8_understanding_direct_to_consumer_genetic_testing_consumers_141208.pdf> accessed 08.07.2016.

[4] The Encode Project Consortium, 'An integrated encyclopedia of DNA elements in the human genome' (2012) 489 *Nature* 57–74.

[5] H.B. Bentzen, 'Persontilpasset medisin – Utvalgte rettslige problemstillinger i tilknytning til klinisk bruk av genomsekvensering og behandling av genetiske opplysninger', forthcoming in *CompLex*.

[6] J.C. Venter et al., 'The Sequence of the Human Genome' (2001) 291 (5507) *Science* 1304–1351; International Human Genome Sequencing Consortium, 'Initial sequencing and analysis of the human genome' (2001) 409 *Nature* 860–921.

[7] International Human Genome Sequencing Consortium, 'Finishing the euchromatic sequence of the human genome' (2004) *Nature* 431 pp. 931–945.

[8] K.A. Wetterstrand, 'DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP)' <https://www.genome.gov/sequencingcostsdata/> accessed 08.07.2016.

In 2014, UK Prime Minister David Cameron announced a £300 million research investment, aimed to map 100,000 human genomes in the UK by 2017, and in time implement personalised medicine as part of routine care in the British health care system.[9] In 2015, US President Barack Obama announced The Precision Medicine Initiative, a research effort on personalised medicine in the US, which was granted $215 million in the President's 2016 budget.[10] Several other countries have also launched or are considering similar initiatives. These initiatives require massive DNA data processing.

2.2. WHY CLOUD?

Relatively recent developments in research have sparked a move to genomics. Gibbons et al. explain:

> Recent genetic research has focused on mapping similarities and differences at the level of the whole genome (that is, all of a person's genes taken collectively). These investigations often use genetic markers called single nucleotide polymorphisms (SNPs) or haplotypes (groups of SNPs that are commonly inherited together). Using these genetic markers makes it possible to screen very large numbers – often many millions – of genetic variations across whole genomes. Scientists, including epidemiologists, thus have begun to investigate correlations (associations) between SNPs or haplotypes and the occurrence of common diseases. Such research investigations study the complexities in the functioning of cells, or the genome, rather than focusing simply on genes. They demonstrate the change from genetic to genomic research. This kind of research, however, requires extremely large biosample collections and associated databases of medical and family history data and environmental and lifestyle information.[11]

In other words, this change in research direction has created an even more pronounced need for DNA data being stored and processed in cloud arrangements. However, there are many reasons why our DNA may end up 'in the clouds'.

International collaborations are necessary and useful in order to achieve progress in the genomic field. Such collaborations can involve the use of human

9 GOV.UK, 'Human Genome: UK to become world number 1 in DNA testing', 01.08.2014 <https://www.gov.uk/government/news/human-genome-uk-to-become-world-number-1-in-dna-testing> accessed 08.07.2016.

10 The White House Office of the Press Secretary, 'Fact Sheet: President Obama's Precision Medicine Initiative', 30.01.2015 <https://www.whitehouse.gov/the-press-office/2015/01/30/fact-sheet-president-obama-s-precision-medicine-initiative> accessed 08.07.2016.

11 S.M.C. Gibbons, J. Kaye, A. Smart, C. Heeney and M. Parker, 'Governing Genetic Databases: Challenges Facing Research Regulation and Practice' (2007) 34 *Journal of Law and Society* 163–189, 166 (internal footnote omitted).

genetic databases. In the absence of a universally agreed definition of genetic databases, we will provide some typical examples.

One example relates to analytical validity. Medical tests are usually evaluated according to the ACCE criteria: *a*nalytical validity, *c*linical validity, *c*linical utility and *e*thical aspects. Analytical validity relates to a test's sensitivity and specificity. A challenge related to genome sequencing is that one can find genetic variants that have not been classified, so it is not possible to determine if the genetic variant is disease causing or harmless. If a harmless variant is classified as disease-causing, this decreases the test's specificity. It is therefore beneficial to establish a database of genetic variants in order to better be able to determine if the variant is normal and harmless or disease causing. Such databases need large numbers, meaning that they ought to be international; in addition, they also need a good representation from the local area of the patient being tested.[12]

Further, researchers and clinicians in the genomics field tend to collaborate internationally, often in large consortia. By centralising data management in the cloud, data and methods can easily be shared by scientists around the world. It is plausible that this is the single strongest reason why DNA data is put in the cloud.

Uploading research data, including research participants' DNA data, into joint databases where other researchers can be granted access is increasingly required by research funders and publishers.[13]

Genomic data requires so much storage space that the most convenient manner for researchers and clinicians to collaborate is through cloud-based processing of the genomic data. It has been calculated that by 2025, human genomes will require 2–40 exabytes of storage capacity.[14] In comparison, YouTube's projected annual storage need is 1–2 exabytes of video by 2025.[15] Shipping hard copies of genome data is not a practical option due to the data size, thus cloud computing is considered the most suitable option for handling such data.

Genomic cloud computing can consequently be defined as Dove et al. do as 'a scalable service where genetic sequence information is stored and processed virtually (in other words, in the "cloud") usually via networked, large-scaled data centres accessible remotely through various clients and platforms over the Internet.'[16] Cloud computing activities are usually divided into three

12 H.B. Bentzen, above n. 5.

13 D.B. Taichman et al., 'Sharing clinical trial data — a proposal from the International Committee of Medical Journal Editors' (2016) 374 *New England Journal of Medicine* 384–386.

14 Z.D. Stephens, S.Y. Lee, F. Faghri, R.H. Campbell, C. Zhai, M.J. Efron, R. Iyer, M.C. Schatz, S. Sinha and G.E. Robinson, 'Big Data: Astronomical or Genomical?' (2015) 13(7) *PLoS Biology*: e1002195 <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002195> accessed 08.07.2016.

15 Ibid.

16 E.S. Dove, Y. Joly and B.M. Knoppers, 'International genomic cloud computing: "mining" the terms of service' in A.S.Y. Cheung and R.H. Weber (eds.), *Privacy and Legal Issues in Cloud Computing*, Edward Elgar Publishing, Cheltenham 2015, pp. 237–259, 240.

categories: Infrastructure as Service (IaaS) which are raw computing resources such as processing power ('compute') and storage, Platform as Service (PaaS) that provide platforms for developing and deploying software applications, or Software as Service (SaaS) which are end-user applications.[17] PaaS genomic cloud computing includes Galaxy, Bionimbus and DNAnexus, and IaaS genomic cloud computing include the Genome Analysis Toolkit.[18] The platforms usually run on clouds provided by cloud service providers such as Amazon.[19]

3. WHY IT IS SO IMPORTANT TO FIND LEGAL SOLUTIONS IN THIS FIELD

Given the obvious benefits society can gain from effective DNA-based research, it may seem somewhat gratuitous to include a section on why we need to find legal solutions that properly regulates the type of situations discussed above. However, it is worthwhile to stop and reflect on the various interest involved. Here we will approach those interest structured around the various relevant actors, including: the researchers who are using DNA databases, the operators of the databases, the individuals whose DNA information is included in the databases, the data subjects' relatives whose information may be revealed, the ethics committees that seek to regulate these databases, as well as a range of third-party users such as law enforcement bodies wishing to access the data in these databases,

The researchers who are using the cloud-based transnational DNA databases have a strong and obvious interest in the legal framework that govern their conduct. As noted by Gibbons et al.: 'If legal standards are unclear and inaccessible, this could … place researchers at risk of criminal or civil liability, and inhibit the progress of research.'[20] Thus, clarity and certainty are two key requirements for the research community, aside from the obvious requirement that the legal framework actually allows for the type of processing the researchers need in order to carry out their research. This same need for clarity and certainty is a key requirement for the database operators.

For individuals, DNA information is one of the most sensitive types of information; as Gibbons et al. remind us:

> It is also worth recalling why research involving human genetics is sometimes considered to be problematic – and, thus, why many believe that genetic databases do

17 W.K. Hon and C. Millard, 'Cloud Technologies and Services' in C. Millard (ed.), *Cloud Computing Law*, Oxford University Press, Oxford 2013, pp. 3–17, 4.

18 E.S. Dove, above n. 17, pp. 240–241.

19 Ibid., pp. 240–243.

20 S.M.C. Gibbons, above n. 12, p. 164.

> warrant special, categorical treatment. It is often claimed that human genetic material is 'special' when compared to other healthrelated materials. The claimed 'special' qualities include that it can be predictive, it is immutable, it is personally identifiable, and it may have implications for others (including family and social groups). These qualities mean that genetic data may have implications for personal life choices, insurance, and employment; raise the spectre of discrimination against individuals or population groups; have significant ramifications for relatives that can shift the balance of rights and interests away from just the individual; contain information which only becomes significant some time after collection; and have cultural significance for certain persons or groups.[21]

In light of this, genetic data can pose a threat to privacy. In this context, it is worth emphasising what can be seen as a definitional mismatch. While data privacy laws typically are aimed at protecting the personal data of identifiable living individuals, DNA information – given the familial nature of genetic information – typically includes information about more than one individual. Data regarding a deceased person in one country may well reveal sensitive information about a relative living in another country.

A special mention must be made of the great potential for secondary use, or misuse, of DNA databases. Even where the data subject is perfectly satisfied with how the research community is handling her DNA data, the database operators may willingly, or unwillingly, share the data for secondary purposes not (specifically or consciously) intended or foreseen at the time of data collection. O'Doherty et al. discuss six such secondary uses:

1. forensic investigations;
2. civil lawsuits;
3. identification of victims of mass casualty events;
4. denial of entry for border security and immigration;
5. making health resource rationing decisions;
6. facilitating human rights abuses and eugenics in autocratic regimes.[22]

While we may perhaps feel comfortable with DNA databases being used for the identification of victims of mass casualty events – such as occurred after the Christmas Day Tsunami of 2004[23] – other secondary uses, for instance civil lawsuits, are more controversial. The noted familial nature of genetic information also makes for complex grey zones in which questions such as whether person A's DNA data also is person B's personal data will arise.

---

21 Ibid., p. 175 (internal footnote omitted).

22 K.C. O'Doherty, E. Christofides, J. Yen, H.B. Bentzen, W. Burke, N. Hallowell, B.A. Koenig and D.J. Willison, 'If you build it they will come: Unintended future uses of organised health data collections', *BMC Medical Ethics*, 2016, 17:54.

23 H.B. Bentzen, above n. 5.

The use of DNA databases by law enforcement agencies (LEAs) and for border control can also be controversial. On the one hand, where a person is aware that the genetic data she provides to a database may be accessed by LEAs, she may be reluctant to provide the sample in the first place which may negatively impact both research and the health of the individual in question. When several people opt not to contribute their samples and data, this can create biases in the research material. On the other hand, LEA access to information held in DNA databases can help solve crime, which is of a general value to society, and a specific value to victims and those wrongly accused. In this context, we can draw parallels to how LEAs have approached data stored by Internet intermediaries such as search engines, cloud storage and social media. We will return to the jurisdictional issues involved below; here it suffices to note that LEAs are displaying an increasing appetite for accessing such data, and that there is a perceivable trend that this appetite is being satisfied by courts approving LEA access to user data held by Internet intermediaries also where the intermediaries are based in other countries and hold that data in other countries.[24]

Further on this, there are secondary uses for which we would never want to see DNA databases being used. Yet we cannot close our eyes to the risk of such databases being misused for discrimination or even ethnic cleansing and genocide. We should always keep in mind the devastating impact the Netherlands population registration system had in the hands of Nazi occupiers in the 1940s. As noted by O'Doherty et al.:

> The death rate among Dutch Jews (73%) was dramatically higher than that among Jews in France (25%) and Belgium (40%), as well as Jewish refugees living in the Netherlands during the Nazi occupation. Seltzer and Anderson argue that this was largely due to the fact that the registration system in the Netherlands facilitated the apprehension of Dutch Jews.[25]

Apart from the research community and the data subjects, we must take account also of the interests of the data subject's relatives and indigenous peoples.[26]

Furthermore, as discussed extensively by Reichel, the role of ethics committees must be considered:

> The main question seems to be how decisions from research ethics committees of different kind may be enacted in composite administrative procedures and allowed

[24] See eg: *Yahoo! v. Belgium*, Belgium Supreme Court decision, Cass. P.13.2082.N., 01.12.2015 and Danish Supreme Court Order delivered 10.05.2012 (Case 129/2011), discussed and analysed by L.B. Langsted and H.L. Guðmundsdóttir, 'Case Translation' (2013) 10 *Digital Evidence and Electronic Signature Law Review* 162–165 <http://journals.sas.ac.uk/deeslr/article/view/2038/1975> accessed 08.07.2016.

[25] K.C. O'Doherty et al., above n. 22.

[26] K.S. Bull, 'Genetiske undersøkelser – Er dagens regulering god nok?' in H. Stenstadvold (ed.), *Georgs bok*, Pax, Oslo 2010, pp. 209–215.

> to have extraterritorial effects. The difficulty lies in the traditional understanding of administrative law of being a legal discipline closely connected to the nation state, with its constitutionally based task to implement the politics of the democratically elected parliaments. Globalisation has challenged this idea and within many areas of administrative law, authorities and public bodies today act beyond the state. The importance of the nation-based democracy as a cradle for legitimate rule making has decreased. … When it comes to administration of ethical approval for medical research, the nation state still seems to remain strong. … [S]everal different administrative jurisdictions, national and European, are involved in one and the same cross-border research project, creating a web of ethical approvals for researchers to adhere to.[27]

As Reichel also points out, Kaye has accurately 'referred to the conceptual underpinnings of current research governance structures as based on the "one researcher, one project, one jurisdiction" model'[28] – a poor fit indeed with the reality of modern cloud-based DNA research.

Finally, one cannot assess the risks associated with the discussed databases without acknowledging the potential for so-called 'function creep'; that is, as correctly stressed by O'Doherty et al., 'shifting social priorities and interests might lead to repurposing of health data collections'.[29] In other words, we can never be sure what the collected data may be used for in the future – a most unsettling thought.

Having noted some of the various key interests involved, we hasten to acknowledge the herculean nature of the task ahead. After all, there are numerous areas of law that impact these databases:

> To get a sense of the sheer range of legal challenges that emerge around genetic databases, it is worth summarizing the principal matters covered by these various governance instruments and common law doctrines. These illustrate the matters which different bodies have seen as requiring attention from regulators. While not exhaustive, in broad terms the following issues feature most prominently: consent; capacity; privacy; confidentiality; the collection, handling, storage, use, and disposal of human tissue and biosamples; data processing, sharing, and preservation; access to data and records by individuals and third parties, including researchers; the use and disclosure of health data and genetic data, including transborder flows; data security and information technology standards; good research practice; healthcare professionals' duties; sharing of genetic information; research governance; ethical scrutiny and ethical approval of research; patenting and other intellectual property rights; ownership, property,

27 J. Reichel, 'Transparency in EU Research Governance? A Case Study on Cross-border Biobanking' in A.S. Lind, J. Reichel and I. Österdahl (eds.), *Information and Law in Transition – Freedom of Speech, the Internet, Privacy and Democracy in the 21st Century*, Liber, Stockholm 2015, pp. 351–382, 376.

28 Ibid., p. 353. See further: J. Kaye, 'From Single Biobanks to International Networks: Developing e-Governance' (2012) 130 *Human Genetics* 377–392, 377.

29 K.C. O'Doherty et al., above n. 22.

and commercial dealings; human rights; benefit-sharing; licensing and inspection of biobanking activities; and the establishment of regulatory authorities, their remits and powers.[30]

If this was not daunting enough, each country will typically have its own laws on these matters with variations between the different countries. And given the cross-border nature of the databases discussed, the operators of such databases, and indeed the users of the databases, are likely to expose themselves to the laws of several countries. Thus, in the cloud arena, the legal issues outlined in the quote above can be multiplied by the (typically large) number of legal systems to which the databases are exposed.

4. ENTERING THE INTERNATIONAL ARENA – PUBLIC, AND PRIVATE, INTERNATIONAL LAW

As already mentioned, the processing of genomic data is regulated differently across the world. To use the European Union as an example; in the EU, data processing, the right to respect for physical and mental integrity, the right to respect for private life, and the prohibition against genetic discrimination, are all considered fundamental rights according to the EU Charter of Fundamental Rights.

One of the changes in the upcoming EU General Data Protection Regulation (EU) 2016/679 (GDPR) as compared to the EU Data Protection Directive 95/46/EC (DPD), is that genetic data is specifically regulated as a special category of data alongside, *inter alia*, health data. Under the DPD, it has been common to consider genetic data health data, which is considered sensitive personal data. The point of departure in Art. 9 GDPR is that processing of genetic data is prohibited. There are exemptions to this, *inter alia*, for scientific research purposes. Nevertheless, there is no doubt that the European Union applies a strict regulatory regime to the processing of genetic data.

EU Member States can according to Art. 9(4) and Recital 53 maintain or introduce further conditions than those set forth in the GDPR with regard to the processing of genetic data. Several European countries have national legislation providing even stricter requirements for the processing of genetic data than those in the DPD and the GDPR.

It is therefore essential to know which laws to comply with and where disputes should be settled. Consequently, we must consider the rules of both private, and public, international law – indeed, to an extent the legal questions that arise challenge the traditional distinction between private, and public, international law.

30 S.M.C. Gibbons, above n. 11, p. 177.

4.1. PUBLIC INTERNATIONAL LAW: THE NOT SO GOLDEN TRIANGLE: SOVEREIGNTY, TERRITORIALITY AND JURISDICTION

Sovereignty usually refers to a state's power and right to govern itself, to make and enforce laws within its borders. It is a descriptive term alluding to supreme power. As Colangelo puts it, the term '[s]overeignty itself offers no analytically independent reason why states have or do not have power; it simply describes the power states do have at any given moment of development of the international legal system.'[31]

Jurisdiction is an aspect of sovereignty, both coextensive and limited by a state's sovereignty.[32] However, a unified definition of 'jurisdiction' does not exist. For the purposes of this chapter, we will use the classical definition provided by Mann as a starting point; 'a State's right under international law to regulate conduct in matters not exclusively of domestic concern.'[33] The 1935 Harvard Research on International Law Draft Convention on Jurisdiction with Respect to Crime ('the Harvard Draft') is, despite it not being a treaty, considered the main framework for assessing public international law jurisdiction.[34]

The territoriality principle is the primary basis for jurisdiction not only in the Harvard Draft, but in international law since the seventeenth century.[35] Under the territoriality principle, a state has jurisdiction over acts that have been committed within its territory. This divides the world into compartments in which each sovereign state has jurisdiction within its borders.[36] Thus, there is a clear connection between sovereignty, territoriality, and jurisdiction.

Territoriality is a thorny concept in relation to cloud computing. Two issues have received particular focus. First, it can be challenging to determine the location of the cloud-based genomic data processing. This question has been subject of much debate. For example, the Article 29 Data Protection Working Party stated that:

> '[C]loud computing is most frequently based on a complete lack of any stable location of data within the cloud provider's network. Data can be in one data centre at 2 pm and

31 A.J. Colangelo, 'Spatial legality' (2012) 107(1) *Northwestern University Law Review* 69–126, 106.
32 F.A. Mann, 'The Doctrine of Jurisdiction in International Law' (1964) 111 *Recueil des Cours* 30.
33 Ibid., p. 9.
34 Draft Convention on Jurisdiction with Respect to Crime (1935) 29 *The American Journal of International Law*, Supplement: Research in International Law, 439–442.
35 C. Ryngaert, *Jurisdiction in International Law*, Oxford University Press, Oxford 2015, pp. 49–100 provides a thorough explanation of the principle and its history.
36 F.A. Mann, above n. 32, p. 30.

> on the other side of the world at 4 pm. The cloud client is therefore rarely in a position to be able to know in real time where the data are located or stored or transferred.[37]

Hon and Millard clarified:

> In most cases, data are usually copied or replicated to different data centres, for business continuity/backup purposes, rather than being 'moved' by being deleted from one data centre and re-created in another. Often the provider will know where a user's data fragments (e.g. for a particular application) are stored, at the data centre if not equipment level.[38]

Second, data can also be located on the territory of states that the data does not have any real substantial connection to.[39] This issue is increasingly moving to the center of the discussion.

The disclosure requests from law enforcement agencies we mentioned above can serve as an example of the challenges territoriality poses for DNA data processing in transnational cloud databases. For identification purposes in criminal cases, law enforcement agencies have shown an interest in obtaining human biological samples from biobanks in order to perform DNA testing on the material.[40] There is reason to believe that such requests will become even more frequent when the material has already been sequenced and it is possible to request access to only the relevant, limited, non-coding parts of the DNA sequence. Thus, disclosure requests from law enforcement agencies to transnational DNA databases should be expected.

Some recent cases illustrate the difficulties that may arise. In the *Yahoo! Belgium* case, the public prosecutor of Dendermonde in Belgium requested that Yahoo! disclose the identity of people who committed Internet fraud via their Yahoo! e-mail addresses. Even though Yahoo! is based in the United States without a branch or offices in Belgium, the Court of Cassation found that such disclosure is not an intervention outside Belgium's territory because Yahoo! has a business link to Belgium.[41] Similarly, a DNA cloud database operator may be based in the

37 Article 29 Data Protection Working Party, Opinion 05/2012 on Cloud Computing, WP 196, adopted 01.07.2012, p. 17.

38 K.W. Hon and C. Millard, 'Data Export in Cloud Computing – How can Personal Data be Transferred outside the EEA? (The Cloud of Unknowing, Part 4)' (04.04.2012), Queen Mary University of London School of Law Cloud Legal Project, p. 7 <http://www.cloudlegal.ccls.qmul.ac.uk/Research/researchpapers/55649.html> accessed 08.07.2016. Cited from C. Kuner, *Transborder Data Flows and Data Privacy Law*, Oxford University Press, Oxford 2013, p. 122.

39 D.J.B. Svantesson, 'A New Jurisprudential Framework for Jurisdiction: Beyond the Harvard Draft' (2015) 109 *American Journal of International Law Unbound* 69 <https://www.asil.org/blogs/new-jurisprudential-framework-jurisdiction-beyond-harvard-draft> accessed 08.07.2016.

40 See for instance the Norwegian Supreme Court decision in Rt. 2006 p. 90 (NOKAS).

41 *Yahoo! v. Belgium*, Belgium Supreme Court decision, Cass. P.13.2082.N.

United States, the cloud may be marketed to and used by Belgian researchers to deposit Belgian citizens' DNA data, and the database operator in the United States may receive a disclosure request from Belgian law enforcement agencies.

In the *Microsoft* warrant case, United States law enforcement wanted information associated with a specified web-based e-mail account stored on Microsoft's servers in Ireland. Microsoft argued that the US enforcement activity is extra-territorial.[42] The United States disagreed, saying that all activities required to retrieve the data can be taken from the US.[43] Both claims are possible.[44] Complying with one country's law can mean breaking another country's law. This places DNA cloud database providers in a precarious position.

To properly engage with the questions at hand, the examples above show that we need to depart from strict territoriality and instead seek alternative mechanisms for delineating rights and responsibilities in relation to DNA data being shared across borders and processed in cloud computing arrangements. We propose the contours of such a solution in section 5 below.

4.2. PRIVATE INTERNATIONAL LAW

As to private international law, questions of jurisdiction and choice of law may arise for several reasons, and the rules of recognition and enforcement may also be actualised in the context of the type of storage and processing of DNA data we discuss. Importantly, both conflicts governed by contract and matters of a non-contractual nature may arise which means that a wide scope of private international law rules need to be considered.

4.2.1. Where disputes should be settled

In the European Union, the Brussels I*bis* Regulation 1215/2012 defines which courts are competent to decide in cross-border litigation between EU Member States in cases concerning civil and commercial matters, such as data privacy.[45]

42 Brief for Appellant, *Microsoft Corporation v. United States* (2d Cir.); for the European Union side, see Brief of *Amicus Curiae* Jan Philipp Albrecht, Member of the European Parliament, *Microsoft Corporation v. United States of America* (2d Cir.).

43 Government's Brief in Support of the Magistrate Judge's Decision to Uphold a Warrant Ordering Microsoft to Disclose Records Within its Custody and Control, *In re A Warrant to Search a Certain E-Mail Account Controlled And Maintained by Microsoft*, 15 F. Supp. 3d 466 (SDNY 2014).

44 D.J.B. Svantesson and F. Gerry, 'Access to extraterritorial evidence: The Microsoft cloud case and beyond' (2015) 31 *Computer Law & Security Review* 478–489.

45 Regulation (EU) No. 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters.

Between EFTA states and EU and EFTA states, the 2007 Lugano Convention, which is almost identical to the Brussels I*bis* Regulation, applies. If the Brussels I*bis* Regulation or the Lugano Convention do not apply, national law applies.

According to Art. 25 Brussels I*bis*, the parties can agree that a court or the courts of a Member State are to have jurisdiction to settle any dispute which have arisen or which may arise. If the parties in the choice of forum clause have chosen a third, non-EU state court, for instance a US court, in principle Brussels I*bis* does not apply.[46]

If nothing has been agreed, Art. 7 Brussels I*bis* provides a default rule in matters related to contract. The courts for the place of performance of the obligation in question are competent. That means that the courts where the cloud services were provided or should have been provided are competent.[47] Moiny illustrates the difficulties in ascertaining where a cloud service is provided or performed, concluding that it could be argued 'that the service is performed where the user normally uses the service, where the service provider supervises and manages the service, or even partially at each place', showing the importance of including an appropriate provision in the contract.[48]

In the United States, three criteria apply: (1) the state must have a long-arm statute; (2) the defendant must have certain minimum contacts with the forum; and (3) the defendant appearing in that forum cannot violate traditional notions of 'fair play and substantial justice'.[49]

Choice of forum clauses are usually part of the cloud terms of service. The choice is often either California or Washington, as both US states are home to many cloud service providers, meaning that non-US users need to be aware of US legislation and legal practices.[50]

4.2.2. Applicable law

Arts. 17 and 2 of the International Covenant on Civil and Political Rights make extraterritorial jurisdictional data privacy claims mandatory.[51] Each signatory

46 J-P. Moiny, 'Cloud and jurisdiction: mind the borders' in A.S.Y. Cheung and R.H. Weber (eds.), *Privacy and Legal Issues in Cloud Computing*, Edward Elgar Publishing, Cheltenham 2015, pp. 118–138, 124.

47 Ibid., p. 125.

48 Ibid., p. 125.

49 T. Mahler (ed.), *Coco Cloud. Confident and Compliant Clouds. First Study of Legal and Regulatory Aspects of Cloud Computing*, p. 129 <http://www.coco-cloud.eu/sites/default/files/cococloud/files/content-files/deliverables/Coco_Deliverable%20D2.2_UO_20141031(1of2).pdf> accessed 08.07.2016.

50 Ibid.

51 D.J.B. Svantesson, 'The Extraterritoriality of EU Data Privacy Law – Its Theoretical Justification and Its Practical Effect on U.S. Businesses' (2014) 53 *Stanford Journal of International Law* 53–102, 77–79.

state is obligated to provide legal protection against unlawful attacks on the privacy of people subject to its jurisdiction and those present within its territory.[52] The convention, however, does not relate to substantive data protection law, such as for instance the particular approach to data protection in the EU.[53]

In the European Union, the Rome I Regulation 593/2008 applies to contractual obligations. According to Art. 2, the Regulation has universal application, so that the law the Regulation specifies should be applied whether or not it is the law of an EU Member State.

The Rome I Regulation's point of departure is freedom of choice. A contract is governed by the law chosen by the parties according to Art. 3. If such a choice has not been made, Art. 4 designates that the law of the country where the service provider has his habitual residence will apply. Irrespective of applicable law, overriding mandatory provisions must be respected. Mandatory provisions are provisions a country regard as crucial for safeguarding its public interest, see Art. 9. The applicable law can according to Art. 21 be refused if it is manifestly incompatible with public policy.

For non-contractual obligations in civil and commercial matters, the Rome II Regulation 864/2007 applies. According to Art. 4, the point of departure is that the law of the country in which the damage occurs is applicable to a non-contractual obligation arising out of a tort/delict. For infringement of intellectual property, Art. 8 states that the law applicable shall be the law of the country for which protection is claimed.

Non-contractual obligations arising out of violations of privacy and right relating to personality, are explicitly exempt from the Rome II Regulation in Art. 1.

For data protection in the EU and EEA, the Data Protection Directive 95/46/EC applies. Article 4 lists three grounds where the national implementation of the Directive applies to the processing of personal data. These are processing in the context of the activities of an establishment of the controller on Member State territory, public international law, and use of equipment on the territory of a Member State. The territorial scope was described in the *Google Spain*[54] case as being 'particularly broad', but in the *Bodil Lindqvist*[55] case the Court of Justice of the European Union had clarified that the Directive should not be interpreted as to be applicable to the entire Internet.[56]

In May 2018, the Directive will be replaced by the General Data Protection Regulation (EU) 2016/679. Article 3 clarifies the territorial scope. In addition to

52 Ibid.
53 Ibid.
54 Case C-131/12, *Google Spain v. AEPD and Mario Costeja Gonzalez*, 13.05.2014.
55 Case C-101/01, *Bodil Lindqvist*, 06.11.2003.
56 C. Kuner, 'Extraterritoriality and regulation of international data transfers in EU data protection law' (2015) 5(4) *International Data Privacy Law* 235–245, 243.

establishment and public international law, the Regulation will also apply to the processing of personal data of data subjects who are in the EU by a controller or processor not in the EU, insofar as the processing is related to offering of goods or services to data subjects in the EU or monitoring of their behaviour in the EU. The territorial scope of the Regulation has far-reaching implications and can be criticised for not distinguishing between the types of data privacy rules that will apply.[57] As suggested elsewhere, a more layered approach where some, but not necessarily all, rules would apply to controllers or processors outside the EU could have been more appropriate.[58] Thus, a transnational DNA cloud database processor in the United States could have been subject to the most relevant, but not all, the EU data protection rules.

5. CONTOURS OF A SOLUTION

We are not here aiming to present a solution to the problems and issues outlined above; our aim is much more humble. All we want to do is to briefly introduce, and bring attention to, some matters that ought to be considered by anyone seeking to propose solutions to the issues we have described and discussed above, thus providing the broad contours of a potential approach going forward. We acknowledge that this is a rather eclectic selection of proposals.

5.1. THE LIMITS OF TERRITORIALITY

As discussed above, territoriality runs as a *fil rouge* through contemporary thinking on jurisdiction. However, its limitations are obvious, not least in a field such as transnational cloud databases for processing of DNA data. In light of this, one important feature of any work towards a solution will be to come up with a better jurisprudential basis for approaching the concept of jurisdiction.

One possibility, previously presented elsewhere, is to look beyond the territoriality principle to the underlying core principles, and adopt the following framework for jurisdiction:

> In the absence of an obligation under international law to exercise jurisdiction, a State may only exercise jurisdiction where:
> (1) there is a substantial connection between the matter and the State seeking to exercise jurisdiction;

57 D.J.B. Svantesson, 'Extraterritoriality and targeting in EU data privacy law: the weak spot undermining the regulation' (2015) 5(4) *International Data Privacy Law* 226–234.

58 D.J.B. Svantesson, 'A "Layered Approach" to the Extraterritoriality of Data Privacy Laws' (2013) 3(4) *International Data Privacy Law* 278–286.

(2) the State seeking to exercise jurisdiction has a legitimate interest in the matter; and
(3) the exercise of jurisdiction is reasonable given the balance between the State's legitimate interests and other interests.[59]

These 'core principles' better correspond to online reality than does the territoriality principle, and adopting these principles as the point of departure for designing rules regarding jurisdiction and applicable law will be more fruitful than clinging on to dated notions of territoriality. For example, this thinking frees from the notion that a country automatically has jurisdiction over any content stored on its territory. These core principles constitute the common core that unites public international law and private international law.[60] Ginsburg has described the approach we here outline as a 'move from an increasingly anachronistic rule, which may be producing more errors than it used to, toward a looser standard that requires careful balancing of interests.'[61]

5.2. HARMONISATION

Greater legal harmonisation internationally may be necessary if the research aims of genomic cloud databases are to be reached. Extensive cross-border scientific collaboration and data sharing requires cross-border legislation. This will also require more collaboration between regulatory authorities. Achieving a minimum data privacy standard can be realistic. As important as this will be, it will not, however, represent a complete solution. After all, even with a minimum data privacy standard, for instance the EU will likely require stricter privacy standards than the minimum requirements for processing of genetic data. For the specific purpose of genomic health care and research, one could therefore also or instead consider an international, preferably global, convention. The problem with a convention is, however, that it may create a too-rigid framework for genome technology that advances far faster than Moore's law. It can also raise questions of whether genomics should be subject to exceptional legal regulation or rather be treated similarly to health data in general. Thus, even where work is undertaken towards harmonisation, other options must be pursued in parallel.

59 See further: D.J.B. Svantesson, above n. 39. This approach to jurisdiction has been endorsed in the Netherlands Presidency of the Council of the EU Debriefing Conference on Jurisdiction in Cyberspace (07–08.03.2016, Amsterdam) doc. 7323/16.

60 D.J.B. Svantesson, above n. 39.

61 T. Ginsburg, 'Introduction to Symposium: Rethinking State Jurisdiction in the Internet Era' (2015) 109 *American Journal of International Law Unbound* 67 <https://www.asil.org/blogs/introduction-symposium-rethinking-state-jurisdiction-internet-era> accessed 08.07.2016.

5.3. BETTER RELATION BETWEEN REGULATION AND TECHNOLOGY

Technology, and the use of technology, are the drivers with which law and regulation try their best to keep up. This is only natural. Nevertheless, it seems to us that, not least in the context of cross-border data transfers, the Global Alliance for Genomics & Health is correct in asserting that: 'In the end, a privacy and security policy must drive the technological choices and final security architecture when sharing data among entities that may span institutional, geographic, and regulatory boundaries.'[62]

In the 1990s, Lex Informatica (or Code) was seen as a viable solution to the jurisdictional dilemmas global technological solutions pose, and it seems to be gaining traction again, now often referred to as algorithmic law. The idea is that a legal regulatory regime lacks the flexibility that the information society requires. Instead, technological rules are used as they do not rely on national borders, allow easy customisation of rules, and benefit from built-in self-enforcement and compliance-monitoring capabilities.[63] The jurisdiction of Lex Informatica is the network itself.[64] Regardless of whether Lex Informatica could be a viable solution to the processing of genomic data in the cloud, we acknowledge that a harmonious balance between technology and legislation is a necessity.

5.4. RISK MITIGATION

Similarly to the situation in other industries, some of the risks of placing DNA data in the cloud may be mitigated by ensuring adequate encryption. As recommended for example by Dove et al., '[r]esearchers should ensure that their own organization has data encryption capabilities and good management infrastructure for control over data stored on a cloud.'[65] Art. 89 GDPR requires that safeguards are in place to ensure respect for the principle of data minimisation in scientific research, mentioning pseudonymisation as one example of such a safeguard.

62 Global Alliance For Genomics & Health, 'Genomic and Clinical Data Sharing Policy Questions with Technology and Security Implications: Consensus Position Statements from the Data Safe Havens Task Team' (18.10.2014) <https://genomicsandhealth.org/files/public/FinalV2_Data%20Safe%20Havens%20Task%20Team_Deliverable_0.pdf> accessed 08.07.2016.

63 J.R. Reidenberg, 'Lex Informatica: The Formulation of Information Policy Rules Through Technology' (1998) 76(3) *Texas Law Review* 553–584, 572.

64 Ibid., 573.

65 E.S. Dove, Y. Joly and A-M. Tassé, 'Genomic cloud computing: legal and ethical points to consider' (2015) 23 *European Journal of Human Genetics* 1271–1278, 1274.

5.5. EDUCATION

In a 2015 determination by the Australian Privacy Commissioner, Timothy Pilgrim, a patient, was awarded a written apology and A$6,500 as a result of the patient's doctor having disclosed personal information about the patient to a law enforcement officer.[66] The officer in question had simply called the doctor and the doctor had disclosed the personal information in question.

Such examples are not rare, and while it may reasonably be presumed that those engaged in research in the DNA field generally have a better understanding of the data privacy considerations involved, one key step forward is to ensure that researchers and doctors gain an even better understanding of their legal obligations. Education in itself is not enough to ensure compliance, but education should be an integrated element of a solution.

5.6. BALANCE OF RESPONSIBILITIES

The role of ethics committees discussed above is, as noted by Reichel, one of two basic points of departure in the area of genomic research. The other is consent:

> In regards to human biological samples in research, two basic points of departure can be identified in national and international law. First, the use of human biological samples in research is conditioned on the informed consent in some form of the donor. Secondly, research on human biological samples should be placed under the review of independent research ethics committees.[67]

We acknowledge the limits to consent, which today is widely used as a basis for depositing individual level genomic data in international clouds. Data subjects (the donors) are often asked to provide their informed consent in relation to matters most lawyers do not fully understand. The impact of choice of law and choice of forum clauses are good examples of this. Such consent can therefore never be truly informed, and to pretend that it is, can only be harmful.

But we must also take care not to create an unworkable 'nanny state' where the individual no longer accepts any personal responsibility. Finding the correct balance of responsibilities is a serious challenge for regulators in this field. Dynamic consent is a promising option for creating better-informed and more understandable consents.[68]

66 *'EZ' and 'EY'* [2015] AlCmr 23 (27.03.2015).

67 J. Reichel, above n. 27, p. 358 (internal footnotes omitted).

68 I.B. Ljøsne, H.J.A. Teare, J. Kaye, S. Beck, H.B. Bentzen, L. Caenazzo, C. Collett, F. D'Abramo, H. Felzmann, T. Finlay, M.K. Javaid, E. Jones, V. Katic, A. Simpson and D. Mascalzoni, 'Dynamic Consent: a potential solution to some of the challenges of modern biomedical research', *BMC Medical Ethics*, 2017, 18:4.

We must ensure that consents are not misused. Inspiration for a workable 'misuse model' may perhaps be found in the 1993 EU Directive on unfair contractual terms in consumer contracts.[69] The Directive provides generally worded provisions meant to ensure that unfair terms are not upheld in consumer contracts. Importantly, those generally worded provisions are combined with detailed examples of types of terms that generally are seen as unfair. This is a highly useful structure that could be replicated in our context here. However, there is another – much less flattering – reason to take note of the 1993 Directive on unfair contractual terms in consumer contracts. One need only sign up for one of the common online services most of us use to come across terms that clearly fall within the unfair terms of the mentioned Directive, and yet they are presented to millions of European users. This tells us something important; in the end, what really matters is not only how we structure regulation in this field, it also matters that the regulation in question actually is enforced.

6. CONCLUDING REMARKS

To date, there has been a paucity of research addressing the international law issues discussed above in the context of DNA data processing in transnational clouds.[70] There is a pressing need for research on this topic. After all, the discussed technologies, and the use of the technologies, are moving forward constantly and rapidly, and the absence of a solid understanding of the legal considerations involved comes with obvious risks.

It could perhaps be suggested that the jurisdictional complexity described above – with a great number of overlapping laws from different countries being applicable to DNA databases – works to promote data privacy. After all, the jurisdictional complexity creates a degree of uncertainty for database operators, and where they wish to ensure compliance with all laws that potentially apply to them, they would rationally abide by the strictest data privacy standards to which they may be exposed. However, first of all, it is questionable that this is how the operators of DNA databases respond in practice. Furthermore, appropriate data privacy protection should be ensured by good data privacy rules, not by unclear rules of jurisdiction. The data privacy protection provided by our complex jurisdictional rules must be improved.

In the above, we have not aimed at being exhaustive on any one topic. Rather, to make the presentation accessible to a diverse audience, we have provided a brief introduction to both DNA data processing and to the legal issues. In addition, we have sought to briefly sketch some of the key considerations as we move forwards towards seeking real solutions in this field. Thus the aim of this chapter is modest, yet important.

[69] Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts.

[70] See, however, the excellent work of researchers such as J. Reichel, above n. 28.