



United Nations  
Educational, Scientific and  
Cultural Organization



**International Bioethics  
Committee (IBC)**

SHS/YES/IBC-24/17/3 REV.2  
Paris, 15 September 2017  
Original: English

## **REPORT OF THE IBC ON BIG DATA AND HEALTH**

Within the framework of its work programme for 2016-2017, the International Bioethics Committee of UNESCO (IBC) decided to address the topic of Big Data and health, including but not limited to the issues of autonomy, consent, data protection, governance, etc.

At the 22<sup>nd</sup> (Ordinary) Session of the IBC in September 2015, the Committee established a Working Group to develop an initial reflection on this topic. The IBC Working Group, using email exchanges, started preparing a text on this reflection between October 2015 and May 2016. It also met in Cologne in May 2016 to refine the structure and content of its text. Based on this work, the IBC Working Group prepared a preliminary draft report which was discussed during its 23<sup>rd</sup> (Ordinary) Session in September 2016. As a follow-up to this discussion, the IBC Working Group started to revise the preliminary draft report between September and December 2016. The IBC Working Group met in Spain in March 2017 to further refine the text. A revised text in the form of a draft report was submitted to the IGBC, the IBC, and COMEST between May and June 2017 for comments. The draft report was then revised based on the comments received. The final draft of the report was further discussed and revised during the 24<sup>th</sup> (Ordinary) Session of the IBC, and was adopted by the Committee on 15 September 2017.

This document does not pretend to be exhaustive and does not necessarily represent the views of the Member States of UNESCO.

## **REPORT OF THE IBC ON BIG DATA AND HEALTH**

### **I. SCOPE AND DEFINITIONS**

### **II. LEGAL REGULATION**

### **III. VISIONS, TRENDS AND CHALLENGES**

### **IV. ETHICAL ASPECTS**

**IV.1.       Autonomy and Consent**

**IV.2.       Privacy and Confidentiality**

**IV.3.       From ownership to custodianship and benefit sharing**

**IV.4.       Justice**

**IV.4.1. Digital gap**

**IV.4.2. Non-discrimination**

**IV.5.       Sustainability with regard to energy and environment**

**IV.6.       Research**

### **V. GOVERNANCE**

### **VI. RECOMMENDATIONS**

### **BIBLIOGRAPHY**

## REPORT OF THE IBC ON BIG DATA AND HEALTH

### I. SCOPE AND DEFINITIONS

1. The digital form of all kinds of data is leading to an exponentially evolving phenomenon called Big Data. It is touching and transforming every area of human life all over the world. In this report the IBC focuses on issues relevant to the area of health on the individual and the public level. It presents recommendations so that the full potential of Big Data can be tapped while at the same time human dignity, fundamental rights and fundamental freedoms are upheld according to Article 3 of the Universal Declaration on Bioethics and Human Rights (2005).

2. Big Data is characterized by the so-called 5 Vs:

- a. *Volume* refers to the huge amount of digital data. It is growing exponentially. While three-fourths of data were analogous in 2000, today more than 99% of all data are digital data. For 2020 there is an estimated amount of 44 zetabyte ( $10^{21}$ ) digital data, for 2025 it is 180 zetabyte (IDC, 2014). But volume of data alone, e.g. coming from whole genome sequencing, does not already constitute Big Data.
- b. *Variety* hints at the fact that there are different kinds of data from diverse kinds of sources. For health care and research several sources are relevant: medical data from individual patient care, public health data, data from different insurances, research data which are collected by researchers and citizen scientists,<sup>1</sup> companies or individuals themselves, lifestyle data e.g. from health apps, data from social networks and data from commerce. These data can be classified in different ways and according to different criteria: there are e.g. personal data, anonymized data, metadata, primary and secondary data.
- c. *Velocity* means the very high speed at which data can be collected and processed. Real-time-tracking and cloud solutions allow for comprehensive processing within seconds, even producing immediate recommendations, e.g. for medication, behavior or nutrition.
- d. *Validity* refers to quality of the data and the question if they really show what they are meant to show regarding content and precision. The context of data plays a major role here.
- e. *Value* finally draws attention to the meaning of the data for a specific question e.g. with regard to a certain disease. Here again it is important to take the context of data into account.

3. Against this background the IBC uses the term Big Data in the area of health as referring to large collections of complex health-related data sets from multiple sources. Typically such data sets cover very large numbers of individuals, but analysis of all available data from one single patient under certain conditions can also be considered Big Data analysis.

4. Big Data cannot be processed with traditional applications, but require enhanced computing power and development of new algorithms. Different kinds of algorithms entail different kinds of ethical challenges. There are, for example, task specific algorithms which are programmed as a defined sequence of clear operations thus allowing for technical transparency. But there are also algorithms in the area of artificial intelligence, machine learning and deep learning where transparency is difficult because algorithms are taught with large datasets and perform operations more or less 'by themselves'.

---

<sup>1</sup> In this report the term citizen science refers to research projects which are conducted by lay people, often cooperating with or being led by professional scientists or research institutions.

5. There is no universal definition of health which could be applied in every context of application to every period in life and in every area of the world. Definitions vary according to perspective (a subjective understanding or objective concept) and according to the purpose of the definition in terms of therapeutic and/or preventive actions which then can be legitimised. For example, the broad definition of the World Health Organization (WHO) is meant as a regulative leading idea to foster global health. It defines health as the state of complete physical, mental and social well-being and not merely the absence of disease or infirmity, thus embracing the whole life of an individual in every respect. When national health care systems have to specify and limit the responsibilities of medical professionals and institutions, a narrower approach is used.

6. The WHO definition shows that everything in life matters with regard to health. In this sense, it confirms a holistic approach to health and blurs the line between health in the medical field and lifestyle. Big Data offers the technical opportunity to support such a holistic view while at the same time giving rise to serious concerns about protection of human rights.

7. In times of information and communication technologies (ICTs) new terms like eHealth (electronic health) and mHealth (mobile health) have emerged. WHO defines eHealth as “the use of information and communication technologies for health” (WHO, n.d.), and mHealth as a subcomponent of eHealth in terms of “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices” (WHO, 2011).

8. There is a day-to-day increasing number of apps (according to estimations there are already more than 300,000) which address health related issues. Regardless of the impossibility to control and to guarantee the quality of all these apps it is hardly possible to sharply delineate medical from nonmedical health apps. They mainly fall into three categories: firstly treatment-related apps, secondly prediction and prevention related apps, and thirdly lifestyle related apps. Accordingly, when addressing Big Data and health in this report, this necessarily entails more than traditional health care and health research.

## II. LEGAL REGULATION

9. There are no specific regulations of the phenomenon of Big Data in the national and international legal frameworks. Nevertheless, there is a complete regulatory framework for personal data protection in many legal jurisdictions, mainly in Europe, of which many rules can be applicable in the area of Big Data, though it is a new reality in the sense of quantity, analysis, accessibility, and application. Furthermore, countries that have no specific laws on personal data protection can still use constitutional and statutory law provisions as well as common law principles for the same purpose. A good example in this regard is most commonwealth countries (UNCTAD, 2016). So there is not a lack of regulation but of specific provisions and perhaps of new principles which are adequate to regulate the new features of Big Data.

10. In the international legal framework, Article 12 of the *Universal Declaration of Human Rights*, adopted by the United Nations (UN) General Assembly in 1948, covers privacy, stating that: “No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks” (UN, 1948). In a similar sense, Article 8(1) of the European Convention on Human Rights states that: “[e]veryone has the right to respect for his private and family life, his home and his correspondence” (CoE, 1950). Article 8(2) adds that: “[t]here shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others” (CoE, 1950).

11. Many of the regulations at the international level have been developed in the context of international data flows mainly due to trade in health services, which leads to cross-border data transfers. The United Nations Conference on Trade and Development (UNCTAD)'s *Data protection regulations and international data flows* (2016) is notable in this regard. Agreement on the core principles can be attributed to the United Nations General Assembly's *Guidelines for the Regulation of Computerized Personal Data Files* (1990), which contain principles to ensure protection of privacy and confidentiality that as a minimum must be provided for in national legislations. These are the principles of purpose-specification and security. The guidelines equally require countries to designate an authority that supervises the observance of these principles, sanction those in breach and prescribe the need to protect privacy during the trans-border movement of personal data. The guidelines were meant to govern computerised and manual files that contain personal information (see paragraph 10 of the guidelines) but the principles can still be applied, to some extent, in the context of Big Data. Other non-legally binding guidelines, which have shaped national legal frameworks, are the World Medical Association's *Declaration of Taipei on Ethical Considerations Regarding Health Databases and Biobanks* (2016) and the *Declaration of Helsinki* (2013). While there is a broad consensus on the core data protection principles at the heart of most national laws and international regimes, the main challenge is divergence in the implementation of these principles as well as in the detailed data protection laws of the world (UNCTAD, 2016).

12. The Organisation for Economic Cooperation and Development (OECD) has developed its *Guidelines Governing the Protection of Privacy and Transborder Flow of Personal Data* (2013). There is also an OECD report which highlights the need for development of appropriate legal frameworks for sharing information (OECD, 2010). This report stresses the need for new legal framework which allows for sharing of health related information between health care professions within and across health care organizations, as well as across organizational and geographical boundaries. The report notes that very few countries in its remit have really addressed these challenges. More recently, in January 2017, the OECD published its *Recommendation on Health Data Governance*. They highlight the importance of complementing legal data protection through education and awareness raising, skills development, and the promotion of technical measures. The Recommendation calls upon countries to develop and implement health data governance frameworks that secure privacy while enabling health data uses that are in the public interest in accordance to twelve high-level principles.

13. With regard to all regulations, an important distinction has to be made between Europe and the US. The European approach is based on a view that privacy is a fundamental human right and it involves top-down regulation and the imposition of across-the-board rules restricting the use of data or requiring explicit consent for that use. The United States, in contrast, employs a sectoral approach that focuses on regulating specific risks of privacy harm in particular contexts, such as health care and credit. This places fewer broad rules on the use of data, allowing industry to be more innovative in its products and services, while also sometimes leaving unregulated potential uses of information that fall between sectors. (USA, 2014)

14. Different approaches in the two regions led to the design of the EU-US and Swiss-US Privacy Shield Frameworks (the privacy shield) by the US Department of Commerce, the European Commission and Swiss Administration for purposes of ensuring compliance with personal data protection requirements in transatlantic commerce (USA, n.d.). The privacy shield replaced the previous Safe Harbor Agreement of 2000 and is subject to annual review to ensure its currency as technology changes and the EU data protection regime is transformed (cf paragraph 27 below) (Weiss and Archick, 2016).

15. The Council of Europe approved in 1981 the *Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data*. The aim was to protect the individual against abuses which may be associated with the collection and processing of personal data. Later, the Council of Europe approved an Additional Protocol to this Convention

(CoE, 2001) regarding supervisory authorities and transborder data flows. It provides for the setting up of national supervisory authorities responsible for ensuring compliance with laws or regulations adopted in pursuance of the Convention. Furthermore data may only be transferred if the recipient State or international organization is able to provide an adequate level of protection. In 2017 the Guidelines of the Council of Europe on the protection of individuals with regard to the processing of personal data in a world of Big Data were set up on the basis of this Convention.

16. The EU has only a limited legal competency on health matters, which can be used mainly to promote cooperation and coordination among Member States. However, in the area of data protection there is a common regulation through Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (EU, 1995). This situation will change in 2018. Instead of a Directive there will be a Regulation which is directly binding in all Member States and is relevant for extraeuropean companies as well (EU, 2016).

- a. The new Regulation provides that “the processing of personal data should be designed to serve mankind. The right to the protection of personal data must be considered in relation to its function in society and be balanced against other fundamental rights. This Regulation respects all fundamental rights, freedoms and principles recognized in the Charter as enshrined in the Treaties, in particular the respect for private and family life, home and communications, the protection of personal data, freedom of thought, conscience and religion, freedom of expression and information, freedom to conduct a business, the right to an effective remedy and to a fair trial, as well as cultural, religious and linguistic diversity” (EU, 2016, paragraph 4).
- b. In relation to consent, the Regulation establishes that it should be given by a clear affirmative act. If the data subject’s consent is to be given following a request by electronic means, the request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided. It later adds that “[c]onsent should not be regarded as freely given if the data subject has no genuine or free choice or is unable to refuse or withdraw consent without detriment” (EU, 2016, paragraph 42).
- c. In the area of research, the Regulation comments that, “provided that recognised ethical standards for scientific research are kept, data subjects should be allowed to give their consent to certain areas of scientific research”, since “the specific purpose of personal data processing can often not be fully identified at the time of data collection” (EU, 2016, paragraph 33).
- d. In the area of public health, the new Regulation mentions that “the processing of special categories of personal data may be necessary for reasons of public interest without consent of the data subject. Such processing should be subject to suitable and specific measures to protect the rights and freedoms of natural persons” (EU, 2016, paragraph 54).

17. The United States of America offer also a legal framework through some recent Acts. There are no regulations that concern Big Data but companies undertaking Big Data processing operations in the area of health need to comply with data protection regulation at the federal level: the Health Insurance Portability and Accountability Act (HIPAA) and the HIPAA Privacy Rule (USA, 1996; USA, 2002). Both regulations require appropriate safeguards to protect the privacy of personal health information, and set limits and conditions on the uses and disclosures that may be made of such information without patient authorization. The rule also gives patients rights over their health information, including rights to examine and obtain a copy of their health records, and to request corrections. The *American Recovery and Reinvestment Act (ARRA)* and the *Health Information Technology for Economic and Clinical Health Act (HITECH)* create significant incentives for an expanded use of

electronic health records and they also contain some rules about data protection, mainly about measures in relation to breaches of health information (USA, 2009).

18. The African Union (AU) and the Asia Pacific Economic Cooperation (APEC) are other examples of regional organizations that have developed regional legal frameworks for the protection of personal data and privacy. The AU's *Convention on Cybersecurity and Personal Data Protection* requires the processing of personal data involving genetic information and health research to be undertaken with the authorisation of the national protection authority (AU, 2014, Article 10[4][a]). The updated *APEC Privacy Framework* provides for "a multilateral mechanism which enables Privacy Enforcement Authorities in the APEC region to cooperate in cross-border privacy enforcement of Privacy Laws" (APEC, 2015).

19. With regard to the protection of data, autonomy and privacy, the legal concept of ownership is important as well. A clear distinction must be made between what constitutes a discovery and what constitutes an invention. Normally, new inventions are patentable, as they entail an inventive activity and are used industrially. On July 3, 2012, the European Court of Justice (ECJ) published its landmark decision in *Used Soft GmbH v Oracle International Corp* (C-128/11). The decision implies that there is specific ownership attributed to intangible goods like software downloaded via the Internet. Although the applicability of this model to other digital goods remains to be considered in future court decisions, the ECJ has opened the door for a discussion on ownership of intangible assets (Hoeren, 2014). Furthermore in an appeal of the judgment in *Football Dataco Ltd v Britten Pools Ltd*, April 2010, the European Court of Justice (ECJ) emphasized that the purpose of the Database Directive is to: "stimulate the creation of data storage and processing systems in order to contribute to the development of an information market [...] and not to protect the creation of materials capable of being collected in a database" (paragraph 34).

20. At the international level the TRIPS agreement, Article 10(2), protects compilations of data or other material in either machine readable or another form. For such data or material to qualify for copyright protection, the Article prescribes that "the selection or arrangement of their contents [must] constitute intellectual creations [...]" (WTO, 1994). The Article further stipulates the scope of rights in such compilations by explicitly providing that "such protection, which shall not extend to the data or material itself, shall be without prejudice to any copyright subsisting in the data or material itself" (WTO, 1994).

21. The Council of the European Union has developed important statements about data base rights (EU, 1995; EU, 1996, Article 7, paragraph 4). For example: specific and separate legal rights last for 15 years, but each time a database is substantially modified, a new set of rights is created. An owner has the right to object to the copying of substantial parts of his database, even if data is extracted and reconstructed piecemeal. The arrangement, selection, and presentation of the data may be protected by copyright, while the database right can protect a whole database.

### III. VISIONS, TRENDS AND CHALLENGES

22. One vision of Big Data in health care is a comprehensive and evidence-based personalized, stratified or so-called precision medicine, which combines the best available scientific knowledge with professional experience of health professionals for the benefit of the individual patient. It is based on technological developments in genomics and other high-throughput 'omics' techniques which have made molecular analysis of human samples orders of magnitude cheaper and more efficient. This molecular data can now be complemented with digital imaging data spanning from the microscopic level to whole body imaging and with environmental and lifestyle information collected from a large number of individuals (e.g. population or patient cohorts), from surveys or from different registries, databases and research infrastructures. Furthermore data can be collected on social environment, communication and behavior, thereby bringing us closer to a more comprehensive understanding of health and disease according to the biopsychosocial concept as put forward

by the WHO. According to this vision, all this data combined with information in the electronic health records (EHR) will in the future provide a fundamentally different approach to diagnose and treat patients in a personalized way, i.e. to offer the right recommendations and individually tailored treatment for a person at the right time. It also can foster patient safety, combining different data from different sources in order to analyze and eventually prevent adverse events.

23. A related vision is that having all this in-depth knowledge of an individual will gradually make it possible to determine his/her predisposition and risk profile to develop a disease and to deliver timely and targeted advice for prevention.

24. Furthermore, it is likely that in the future health care will move more and more towards remote collecting data for diagnosing, monitoring and supporting therapy. Collected data can then be used to detect early warnings of disease, like an imminent heart attack, and make recommendations for adequate behavior. It can also contribute to improved telemedical health care for people living in remote regions, allowing better access to quality health care and thereby contributing to global health.

25. Patients can have the possibility of greater access to and control over their data, e.g. having their EHRs in their smartphone. The smartphone could actually turn out to be a central device for coordinating one's own health care and for creating one's own health network, thus fostering autonomy and health literacy.

26. Having better access to health information as well as individualized profiling and recommendations – provided they are quality approved – citizens will potentially profit from better understanding their health status and health improving behavior. However, there is a human-dependent challenge: the fact that even if one understood the information it does not necessarily mean acting accordingly, and in fact, often no health-promoting action whatsoever is taken. Tobacco smoking is a striking example.

27. Health apps are playing an increasing role in health care and research. Major improvements in several areas are hoped for, e.g. in monitoring of health related measures, positive health behavior changes, and diagnostic support. However, so far, their quality is not officially assessed and monitored, and there is hardly any specific regulatory framework or meaningful vigilance system in place. Frequent updates are an additional obstacle for valid assessment of health apps.

28. Against this background, at least four paradigm shifts in individual health care are likely to occur: a shift from disease orientation to health orientation; from focus on therapy to prevention; from health to lifestyle counseling; and from the role of a patient to the role of a user, customer or digital citizen.

29. For the pharmaceutical and medical device industry there is the hope that Big Data will foster the understanding of diseases and their underlying mechanism, thereby leading to the development of new targeted drugs, devices and treatments. Such Big Data is also expected to help design stratified clinical studies, thus reducing the number of participants and costs, and give quicker results.

30. But these visions of Big Data do not only pertain to individual health care and research. Big Data approaches are also supposed to provide a lot of new information in order to strengthen the evidence base for public health policies, e.g. enabling better risk-adjusted prevention strategies for defined target groups.

31. Regulators might better understand and control study designs and their policies might benefit from improved pharmacovigilance. Once a new drug has entered the market, Big Data allows for real world data collection and assessment in a large number of patients over a long period of time. As the majority of the global population lives in areas covered by mobile cellular networks, the number of citizens who can contribute data is bound to increase.



32. Furthermore, Big Data can contribute to support learning health care systems (IOM, 2007; IOM, 2013). Everyday experiences can inform the best and most efficient way to administer diagnostic, therapeutic and preventive measures, as well as shaping structures for health care and research. This affords analysis of real world data from health care in a structured and quality controlled way. Examples are the work towards learning health care systems of the OECD, EU projects such as the Coordinated Research Infrastructures Building Enduring Life-science Services (CORBEL) Project (<http://www.corbel-project.eu/about-corbel.html>), and practical instruments that facilitate data sharing such as Innovative Medicines Initiative (IMI), and industry initiatives (OECD, 2013b; OECD, 2015; IMI, 2014; PhRMA, 2014). Recently, the European Medicines Agency (EMA) published its guidance on the release of anonymised individual patient (trial) data (EMA, 2016).

33. Having in mind the broad definition of health by the WHO and recognizing that the major part of an individual's health status does not depend only on health care but on social determinants, including, but not limited to, social inequalities, education, lifestyle, and environment, Big Data opens up the way to a holistic view on health by bringing together different kinds of data e.g. from registries, apps and health records.

34. These visions and trends are accompanied by major challenges which are of a technological as well as of an ethical, social and legal nature (see Chapter IV).

35. Welcoming the possibility of a holistic view on health, the line between the health care sector and other societal sectors is increasingly blurring. In addition to the traditional sources of health data, such as medical records and laboratory results, other sources which are not traditionally regarded as health related are used, for example social networks or consumer data, public data sources from non-health areas such as the Internal Revenue Service, Education or Social Services Departments. Search engine providers also request and collect much information on their users, which they later process and sell to different companies which in turn use 'personalized marketing strategies', offering users different promotions based on their search histories or participation in online groups, for instance. Such customer profiling also includes health issues – a simple 'private' internet search on a personal or familial condition becomes publicly accessible information. Therefore, there are obviously complex challenges to data protection and privacy as well as to the quality of data.

36. Machine-learning prediction methods have been very productive in medicine. However, it is essential to understand underlying assumptions and ensure that conditions like stability (namely the conditions in which the data were collected remain the same) are met to secure the quality of data as well as the validity and usefulness of the conclusions. Big Data mainly allows for finding patterns and correlation, which can be misunderstood as causal relationships so that inappropriate and harmful consequences can emerge.

37. Furthermore, the algorithms used in pattern recognition and health applications can entail biases from the beginning, or can be hardly understood and controlled in terms of technical transparency.

38. One problem is the possible inaccuracy and imprecision in the collection and classification of data because of the enormous quantity, heterogeneity, fast accumulation, fast management and different contexts of data. There may be errors in the process of data collection by the patient/citizen, by the medical professional (e.g. incomplete registration, inadequate classification, imprecise analysis) or by the (bio)informatic technology (due to e.g. biased or imprecise algorithms, lack of interaction between clinicians and computer scientists). Inaccuracy and mistakes in the collection and analysis may have adverse implications both for the single patient/citizen and for society (EC, 2015).

39. So, appropriate skills and competences by all involved actors are a major challenge as well. Health professionals will have to understand and to properly verify and apply available data. Researchers, citizens and policy makers have to understand the relevance of different contexts of data and the implications of algorithms. This requires transparency and

interpretability of data and their contexts. This is only possible if algorithms are accessible, and if experimental design and optimal resource allocation policies are given.

40. The line between different professional disciplines will also blur. At least some diseases will no longer be understood according to the affected organ, but rather according to the underlying molecular mechanism, the pattern of mutations and variants in the 'omics'. A multidisciplinary approach for health care and research, as is already implemented in oncology in many countries, is needed to optimize the treatment of individual patients which may also result in new professional disciplines. Furthermore the integration of disciplines like social science, behavioral economics, epidemiology etc. from the beginning seems to be adequate in view of a comprehensive understanding of health as well as of the use of data in causal inference methods.

41. Big Data developments offer a lot of opportunities for improvement of health and could mean saving money for keeping sustainable health care systems. However, these technological developments also present some challenges for health professionals. The implementation of many technological procedures can reduce the number of at least some kinds of health professionals that the health care system need, they will probably need different skills, and it might cause a gap between different generations of health professionals. Therefore, it's crucial to adapt the educational system to this new context and improve the continuous training of health care professionals.

42. Eventually Big Data and an environment full of sensors lead to the effect that individuals become more and more transparent while the technological environment, algorithms, as well as the consequences of data analysis and the underlying weighting of factors become increasingly opaque. This leads to several ethical challenges especially with regard to autonomy, privacy and justice (see Chapter IV).

43. In many countries, particularly in the developing world the ethical and scientific infrastructure for research are still developing and many are still far from harnessing the potential benefits that the use of Big Data promises. Access to data remains a challenge and the health information platforms that can support the storage of huge amounts of data and to translate existing data into real action are still lacking. This calls for some attention to capacity building. Capacity strengthening is an important way of ensuring that developing world scientists and health professionals do not become just data collectors in international collaborative projects but active participants in transforming data into tangible health benefits for their own populations.

44. Furthermore, advanced health care towards precision medicine is expensive and limited to richer countries and Big Data approaches do not yet adequately cover diseases affecting people in countries with less advanced health infrastructures. To implement Big Data-driven precision medicine new solutions for data handling in actual health care settings are needed. Only then could Big Data be used to improve medical interventions and health services, prediction and prevention strategies and health policies in general. This is especially a concern for developing countries with limited financial and human resources, and different sets of health priorities.

45. It is very difficult to foresee how fast and to what degree the trends described above will be made a widespread reality. Despite the sketched optimistic vision, it is important to be cautious and to avoid exaggeration of the current state of scientific knowledge and the potential benefits of Big Data and precision medicine for health care globally. On the one hand, the 'hype of Big Data' can lead to overstatements and unrealistic estimations, as well as to marketing of unproven products and services. Furthermore it can lead to unbalanced health care policy priorities which can have especially harmful effects for countries where access to essential services is not guaranteed. On the other hand, there is the danger of losing the benefits of Big Data by neglecting its potentials. A balanced way of dealing with hopes and promises is very important.

## IV. ETHICAL ASPECTS

### IV.1. Autonomy and Consent

46. Autonomy of an individual in terms of exercise of his or her self-determination includes seven dimensions (Mertz et al., 2016):

- a. *Competence* of the individual concerned to access, to understand, to assess and to apply relevant information;
- b. *Information* has to be available that is relevant and understandable for the question at stake;
- c. There has to be a *choice* between different options, be it the one between doing or abstaining from doing something or the one between doing different things;
- d. *Values* of the individual, his or her preferences and attitudes are taken into account in deciding and acting;
- e. *Voluntariness* is granted so that the individual can decide and act without inner or outer coercion;
- f. *Formation of will* refers to the ability of the individual to choose an aim and appropriate means to reach it;
- g. *Action* can mean a conscious doing or refraining.

47. The main guarantee, which has traditionally been established to protect the autonomy of human subjects in health care and research, is consent. Autonomy and responsibility, as well as consent and protection of persons without the capacity to consent, are addressed in Articles 5, 6 and 7 of the *Universal Declaration on Bioethics and Human Rights* (UDBHR). Article 6 provides that any preventive, diagnostic and therapeutic medical intervention as well as scientific research should only be carried out “with the prior, free, express and informed consent of the person concerned” (UNESCO, 2005). The *Nuremberg Code* concentrates on consent and is at the origin of the concept that participation in research is a voluntary activity. The *Declaration of Helsinki* also enshrines consent as a main guarantee.

48. The scale and complexity of personal data processing in the area of Big Data make it virtually impossible for people to keep track or make meaningful decisions about all uses of their personal data. There are major problems with regard to all seven dimensions which are mentioned above. Even if individuals could keep track and would be informed, their decisions might be skewed by various decision-making difficulties. This leads to a loss of autonomy which in consequence means a loss of control and a loss of freedom from decisions which are made by the technological environment and automated processes. Governance has to compensate for this loss of control, ensuring that all the dimensions of autonomy are respected and supported (see Chapter V).

49. A further problem of informed consent given electronically (informatic consent) is the impossibility to identify the age of the user, and so to tailor the information to be presented according to his or her age and capability of understanding (EC, 2012). The education of minors as active users of these new technologies is particularly important.

50. In the context of clinical trials, consent has traditionally been specific, that is, subjects are asked to consent to a specific research project. In many such cases, the consent obtained from research participants does not extend to the use of samples or data outside the original or primary research they have consented to. Valid consent requires that adequate information is disclosed and that potential subjects understand the nature, risks and potential benefits of the research and thus grant voluntary informed consent.

51. In the area of large databases like biobanks, and even more in the area of Big Data research, the potential secondary use of data in the future cannot be predicted. It also involves interrelationships between multiple and changing data sources, both medical and non-medical. This has led to calls for a new, more appropriate model of consent that allows a wide

use of data while at the same time respecting the participant's or patient's autonomy. Already the field of biobanks has put forward new suggestions for consent provision such as what is called *broad consent*.

52. Broad consent can cover a wide range of activities while at the same time being narrowed down to limited, though just vaguely specified uses such as research into the causes of complex diseases. Thus broad consent is not the opposite of 'specific' consent because according to the understanding of the IBC it is not blanket or open consent, which does not have any conditions attached. Broad consent typically operates at a higher level of abstraction. It leaves the essential informed consent model intact, but an individual simply consents to a *range of* possible research that could be done with her/his information in relation to a specific area or line of investigation. The conditions for broad consent to be valid is the need for some form of governance and safeguard where, for example, a relevant committee reviews a proposal to ensure that the rights and interests of individuals are adequately protected or where safeguards mandated by legislations are in place.

53. This formula of broad (informed) consent is getting more and more accepted (WMA, CIOMS), and is utilized in Europe, where individuals broadly consent to all possible research with their tissue related to the same or similar area or line of research. It is also a model that is gaining attention in developing countries, e.g. in the context of genomics and biobanking, but there are also concerns because of some negative experiences in the field of biobanks. With appropriate safeguards, such as controls to ensure that the data will not be used to make a decision about the individual or otherwise used in a manner that will impact the individual and/or the community, the use of broad consent is suitable for research purposes that contribute to public good.

54. Another approach is opt-out consent according to which health data can be used for research purposes unless an individual affirmatively opts out. This approach requires adequate information beforehand, addressing the potential implications of permitting biospecimens and personal data to be collected. So it requires practitioners to provide adequate education and information although time often is scarce. Differing from the research setting the virtual impossibility to opt-out from the commercial health-related databases (of mobile phone and internet operators, apps, commercial and consumer behavior databases) is a particular concern, as there is virtually no control over such databases.

55. According to the principles of participation and transparency a further model which depends on the availability and accessibility of ICTs has been proposed. It is called 'dynamic consent'. Based on an initial consent from the participant it entails an update on the use of data in a continuous way so that the individual can opt out for specific uses of data while allowing the use for other purposes. It requires the involvement of public powers both as a guarantee of the individual rights as well as the promotion of individuals' participation through education and information. It calls for empowering subjects and patients to be able to monitor the use of their health data. Dynamic consent could in some settings allow for control of data access by individuals, enabled by mechanisms such as consent portals. Continuing participation can have the advantage of allowing participants to shape the possibilities of research through their decisions about what uses of data to permit by effectively 'voting' for those uses by consenting to them. The individual effectively participates in the enterprise (Nuffield Council, 2015). In a way, the project becomes a shared or joint enterprise of the patient along with the researcher, and so the individual is not being used or exploited. This model of consent could work well in developed countries with advanced technologies that allow patients to monitor how their data is being used through a database system, and people who have the time and willingness to devote it to follow the use made of their data. However, its implementation in less developed settings could face several challenges such as the lack of local technologies and low literacy rates, the need to devote people's time to other activities, or the lack of public awareness and education so that patients' comprehension in the context of health care and participation in research could be hindered.

56. Consent is a major challenge with regard to health apps and in the area of social media and networks. The digital information is generally written in small characters to be displayed on smartphones, sometimes with no real alternative of dissent. The process of giving informed consent on screen (and not on paper), in a setting where the digital user does not have access to consultation, often leads to clicking in an immediate and fast way without sufficient time for making an informed choice and without the possibility of ascertaining actual awareness, competence and voluntariness. Trade terms, which are usually too long and for most people hardly understandable, are seldom read at all. In this sense it is doubtful that the 'one click, accept all'-procedure is really a free and informed consent. Electronic means of conducting research and health surveys may however be efficient and effective as long as there are safeguards implemented to ensure that the participants' autonomy is respected.

57. Facing the challenges for obtaining an autonomous consent for the collection and use of Big Data and the limitations of all proposed models of consent, it is important that autonomy is supported by public education which allows users, patients and research participants to understand the range and the impact of possible uses of their data.

58. Though it must be stressed that autonomy links directly with a person's dignity, there are also major collective interests which support research being carried out in the field of Big Data and health in order to promote evidence-based precision medicine as well as learning health care systems. Therefore, a balance between individual and collective interests must be sought. Thus a solidary claim to consent to participation in databased health research by contributing one's anonymized – not personal – data for large databases can be discussed. According to the principle of solidary reciprocity, an individual who profits from former health research can be supposed to have an interest that other people can profit from progress in research as well – at least when everyone can be sure that the outcome is beneficial in terms of a democratically defined public good, and when misuse and unlegitimated deanonymisation of data are strongly prosecuted and punished. In such a framework, there are at least two main elements which could be useful to balance the individual and collective interest: firstly the vulnerability of the subjects from whom data are collected; secondly, the purposes for which the data will be used.

59. In case research is intended that falls outside the range of the broad consent that was obtained for the use of this data, specific consent is necessary for secondary data processing. This is an essential principle to guarantee confidentiality and data privacy. However, secondary analysis of data could be ethically admissible without a new informed consent for such secondary use provided that all the following requirements are met:

- a. appropriate legal foundation;
- b. evaluation by the Research Ethics Committee (REC);
- c. adequate technical procedures in order to prevent researchers and third parties from accessing personal data, such as pseudo-anonymisation;
- d. overriding public interest in this health research;
- e. infeasible to obtain a new consent;
- f. data must have been collected according to ethical and legal requirements.

## **IV.2. Privacy and Confidentiality**

60. According to Article 9 of the UDBHR, the privacy of the person concerned and the confidentiality of their personal information should be respected (UNESCO, 2005). Such information should not be used or disclosed for purposes other than those for which it was collected or consented to. In times of Big Data, such protection of privacy and confidentiality is facing several challenges.

61. Traditional data protection principles such as purpose limitation, data scarcity and minimization, special protection of sensitive data, fair processing, and safeguarding the rights of the persons concerned have been widely implemented in order to protect privacy, though

there are different protection schemes and underlying concepts in different regions of the world. By way of contrast, Big Data inherently entails change and openness of purpose, limitlessness of data, intransparent processing, little protection of data in order to gain as much knowledge as possible, and intransparency for persons concerned. This is especially true with regard to algorithms becoming more and more complicated and incomprehensible in times of artificial intelligence and deep learning.

62. Furthermore, anonymisation of personal data in certain circumstances, and depending on the accessibility of other sources of data, no longer provides sufficient protection of privacy and confidentiality. Integrating large amounts of data from different kinds of sources may make the reidentification of the person concerned possible. For anonymisation to remain a useful tool for the protection of the individual, and at the same time offering the possibility of using Big Data for the public good, safeguards which mitigate the risk of re-identification are important.

63. In addition more and more differentiated profiling is done, which allows for acting on groups according to specific profiles (the so called 'group identity'). Thus privacy of the individual is no longer adequately protected even if the individual's name is not known, but e.g. the IP address of his or her smartphone or computer. Thus group privacy is an increasingly pertinent concept in Big Data applications. The Parliamentary Assembly of the Council of Europe even proposed the recognition of a new right of a subject to not be profiled (Rathenau Instituut, 2017).

64. Such a new right acknowledges that in times of Big Data, privacy as a concept is more than mere data protection. When collecting, storing, processing, analyzing, and applying data which then are collected, stored etc. again and again (the so called cybernetic loop) it also encompasses the intrusion into the private sphere by mostly commercially-driven information, offers, advertisement and the like on the screen of electronic devices according to one's personal profile which is constantly assorted by intransparent algorithms on the basis of a variety of data from diverse sources. This intrusion touches several dimensions of private life like personal communication, location, and participation in associations. Against this background, the right not to be profiled seems even more important.

65. The right to privacy is tightly linked to the right to freedom with its diverse legal and ethical aspects like freedom of speech, of association, of location, of movement and space, of beliefs, thoughts and feelings, and of behavior. Against this background, the IBC uses the term privacy in the sense of a right to respect for private life in relation to those areas of life or those data that individuals want to keep reserved for themselves or, at least, for some specific members of their families or relationships.

66. While individuals have privacy interests, they also share public interests in the wider use of data for health research. This broader public interest may come into conflict with individual privacy. But the relationship between privacy and public interest is not simply one of opposition. The two are mutually implicated: there are private interests in the achievement of common goals and a public interest in the protection of privacy that encourages cooperation. This complex relationship leads to a need to reconcile the articulation of the private within the public and the public within the private.

67. From several surveys we know that people are aware of having little control over collection and mining of their data although a majority of them wants to have a certain degree of control. In the long run this might lead to a lack of trust. The loss of trust might well cause serious damage to future essential endeavours and projects to foster public health. In order to protect privacy and the public good at the same time, it is essential that the population is able to trust in the good use and protection of their health data through a mix of approaches like law, governance, public surveillance, privacy by design of a device or an app and privacy by default e.g. with regard to pre-installed consent prerequisites.

68. New models of participation in data collection and mining, as outlined in the chapter

on autonomy, proper public education on the implications of the use of Big Data, effective data protection management, accountability of use to both participants and society in general, as well as innovative models of data ownership and trusteeship are to be developed in order to protect privacy in the broad understanding outlined here.

### **IV.3. From ownership to custodianship and benefit sharing**

69. Ownership of personal data and the rights entailed, among which we might highlight the restriction of access by third parties, must be recognized, without any doubt, as being of the person/group from whom the data derived. Ownership entails the right of data subjects to track and check the existence and manipulation of their personal data. Maintaining control of what we are, what we do and what we think and what we present to others and society in terms of personal information is a key facet of liberty in the 21st Century (see also Chapter IV.2.).

70. In the area of Big Data, it is important to distinguish between two different concepts linked to the idea of ownership, as it is done in similar areas like biomedicine: that of ownership of personal data and ownership of operational results, ownership as a mechanism to control data and ownership of products like algorithms, drugs or monitoring tools etc., and intellectual property.

71. It is argued that the benefits of data will be harnessed through fair data sharing practices with the wider scientific community and that this process will facilitate scientific discovery. There is also growing scholarship on the appropriate ethical frameworks that should guide best practices in data sharing in international research collaborations. This includes the use of deliberative methods to solicit the views of key stakeholders. There seems to be a growing consensus in the scientific community on the importance of sharing research data more broadly. At the same time, there are serious concerns about data sharing, including consent, ownership of data, related intellectual property rights and privacy issues.

72. In this context, Big Data phenomena like real time analysis, linking and sharing of large databases, and use of databases for different purposes in health care and research at the same time and again for different purposes at a later time, make it increasingly problematic and unfeasible to solve the ethical and legal challenges by concepts of ownership. Due to technical equipment, skills and financial resources, only few companies and institutions can deal with very large databases, thus excluding others from gaining knowledge and developing beneficial tools for the health of all. Thus traditional concepts of ownership seem no longer an adequate normative framework especially with regard to the so-called digital divide. The phenomenon of Big Data rather raises the need for developing an alternative normative framework that embodies new concepts to balance legitimate interests and benefits.

73. The new context can be seen as an opportunity to revisit our traditional vision and develop new ethical and legal ways for a real scheme of sharing benefits. It is necessary to reconcile all the rights and interests which overlap in this field, such as those of the person from whom the data derives, those of the researcher, those of the companies and organizations who use the data, and those of society in general who may benefit from such use. This is meant to shift the focus from ownership to custodianship and the responsibilities of all stakeholders with benefit sharing being a crucial part. Custodianship means the responsibility for the safety and well-being of someone or something and represents ethical values like care, custody (medical ethics), protection and trust to the guardianship or the safekeeping.

74. Against this background, Big Data can be framed as a common good of humankind. Advancements and new opportunities provided by science and technology might help reduce and not deepen the inequalities that prevent many human beings from enjoying the highest attainable standard of health, both at the national and the international level. This is in harmony with Article 2 of the UDBHR which states: "(f) to promote equitable access to medical, scientific and technological developments as well as the greatest possible flow and the rapid sharing of

knowledge concerning those developments and the sharing of benefits, with particular attention to the needs of developing countries; and (g) to safeguard and promote the interests of the present and future generations” (UNESCO, 2005).

75. The future challenge will be to create the best data infrastructure available for use by anyone in the world, developing new ways to get economic profits through parallel or complementary services or donations. There is the need to improve solidarity not only from the citizens, allowing them to share or donate their data for the common good, but also for companies and private actors to share their work for the same end. Some projects from several governments, NGOs and private companies under the common name of Open Data and Open Science are good examples of the near future, where “the best data infrastructure [could] be made available for use by anyone in the world, a powerful platform for experimentation, discovery, and the creation of new and better ways of living” (Nielsen, 2014).

76. The concepts of responsibility and common good reinforce what the IBC has advocated for in a previous Report on the Principle of the Sharing of Benefits: “[b]y accepting the idea of universal human rights, we acknowledge that these advancements cannot be a privilege. Progress cannot deepen the existing faults of inequality between peoples and countries. At the same time, the report acknowledges that solidarity through participation and not beneficence is the bond of sharing that needs to be boosted. This is the only way to bring together development and respect for all” (UNESCO, 2015). This includes the concepts of capacity building and science education, brain circulation, open access to health-related information, and empowerment and participation in the production of knowledge.

77. Several programmes have been initiated to strengthen capacities of researchers and research institutions across the globe, especially in developing countries. In the context of genomics and biobanking research in Africa, for example, the Human Heredity and Health in Africa Initiative (H3Africa) aims to strengthen genomic and biobanking research expertise on the continent to address health problems in both communicable and non communicable diseases (H3Africa Consortium, 2014). The expectation is that this initiative will enable African scientists to contribute meaningfully to cutting edge scientific projects and also become leaders in the field of genomics, bioinformatics and data science. Ultimately, the aim is to improve the health of African people and the global community.

78. Developing an ethical framework for benefit sharing in the use of Big Data raises the question on how it will be applied in different settings. A lot will depend on the context in which the data is being used (in research, health care or lifestyle-settings), the nature of the benefits, and which individuals or groups have a legitimate interest in the use of the data. This diversity, as well as the balancing of the advantages and risks of specific frameworks for benefit sharing, can most comprehensively be addressed by a multi-tiered governance model (see Chapter V). It may, for example, entail giving a considerable amount of a data based financial profit for the benefit of the group who substantially contributed their data to make insights and products possible.

#### **IV.4. Justice**

##### ***IV.4.1. Digital Gap***

79. Big Data needs tools and providers to access and manage them. The digital gap (called digital divide as well) constitutes one of the major challenges in the process of democratizing information and communication, and thus development. Digital technology is, in fact, a major asset for development and for fighting poverty so that it is one of the objectives of the Millennium Development Goals of the United Nations regarding the strengthening of the global partnership for development. As a result, the gap is diminishing. One target was to ensure that the benefits of new technologies, and especially ICTs, were made available for everyone. This program actually increased the level of information of the populations, in particular the poorest and the most isolated. Regarding the health sector, this strategy strengthened the prevention of health problems thanks to a better health promotion, but also



significantly expanded the supply of care through telemedicine in particular. This is a positive development regarding people's awareness and knowledge, which is the basis of any public health action. The good performances in mobile telephony contributed to the implementation of initiatives in order to strengthen the knowledge and skills for the public and actors in the fight against non-communicable diseases. For example, the *Be Healthy Be Mobile* program is implemented in several countries (e.g. India, Egypt, Senegal, Costa Rica, Zambia).

80. Internet coverage increased, from just over 6% of the global population in 2000 to 48% in 2017, and 53.6% of the households covered worldwide (ITU, 2017). Therefore, 3.6 billion people are connected to a global network of apps and contents. However, in many low-income countries, limited international bandwidth and poor national infrastructure hinder the provision of broadband Internet services at affordable prices. These limitations have concrete effects on the types of services and applications users can access. Consequently, the developing countries and the least developed countries experience limitations to integrate the management of Big Data in their offers. In addition, the average price of services remains relatively high in many of the poorest countries where the maintenance and up-dating of the tools (i.e. cell phone) might be difficult.

81. Despite all these advancements in access to ICTs, huge disparities remain and tend to grow between countries, but also within countries. The analysis of Internet household coverage reveals clear differences according to the country's economic level. Thus, in 2017, Europe has coverage of 84.2%, against coverage of 18% for Africa (ITU, 2017). Less than the half (41.3%) of the population of developing countries uses the Internet, against 81% in developed countries (ITU, 2017). Consequently, although problems of isolation and accessibility to health facilities and professionals are given priority in most developing countries, technological innovations such as tele-expertise, tele-diagnosis or tele-consultation, constituting effective solutions are not yet sufficiently implemented due to lack of resources.

82. There are also significant inequalities among countries in terms of ICT skills, digital literacy, and the existence of relevant local content. So even if the innovative content is offered through ICTs in developing countries, they cannot be implemented there because it requires infrastructure and human resources for their adaptation and evaluation. This situation raises ethical problems such as massive transfer of data without clear conditions, coming with the risk of 'data-piracy', or inadequate governance policies (Wyber et al., 2015). Indeed, masses of information are transferred from developing countries with limited human and technical resources, to be processed without some assurance on the protection of data transferred.

#### **IV.4.2. Non-discrimination**

83. In many countries access to health care is not based on a public model but on private models, which are managed by private insurance companies with varying degrees of involvement, control and monitoring by the public authorities. In consequence, subjects can be denied or impeded from accessing insurance due to disclosure of health data. But should insurance companies become aware of the specific risks of a person in detail, it is very difficult to speak of an insurance model which has traditionally been based on balancing of collective risks rather than detailed individual risk profiles.

84. Using ICTs and Big Data approaches, behavior-based insurance models are coming up in different types of insurance, including health insurance. Persons are offered reduced insurance rates or other kinds of premiums if they transfer their data, e.g. about sportive activities, nutrition habits and the like, on a regular basis via tracking, wearables and apps. This can turn out to discriminate against three groups of people either in the form of not having access to benefits or, what is to be expected in the long run, in the form of suffering from disadvantages. Firstly, those people who do not want to share their data with the insurance company are discriminated against, regardless whether they can fulfill the norms like walking 10,000 steps per day or not. Secondly, those people who cannot fulfill the norms because of a disability, a disease or a great misfortune are affected, regardless of their willingness to share the data. But exactly these people are in special need of solidarity and social support.

Thirdly, those people who have different beliefs than the insurance companies about what health and a healthy life is about will be discriminated against. The authority to define which health related behavior is rewarded – e.g. walking 10,000 steps but not meditating every morning – should not be with insurance companies. It is rather a scientific and social issue.

85. There is also a risk for discrimination coming from Big Data based breaches in anonymisation. Aggregating data opens the possibility of re-identification through cross-referencing with data concerning ethnic background, locational data, age, other metadata, health records or genetic data (Choudhury et al., 2014).

86. In addition, there are powerful algorithms, which can profile individuals without having personal data about them in order to form groups according to geographical, socioeconomic, ethnic or other characteristics. The anonymisation of personal data matters little then if the outcomes affect the groups to which they belong with a risk of discrimination and stigmatisation. Such effects impact also those members of the community, who did not give their consent and who didn't contribute their data but nevertheless are members of the profile-defined group.

87. Furthermore, a loss of freedom can result from a subtle pressure to adjust behavior. Predictions are constructed on the frequency of behavior. The consequence is on the one hand the stigmatization of infrequent behavior (of minorities), and on the other hand a hidden coercive pressure to adopt 'frequent behaviour' in order not to be excluded. This pertains to groups as well.

88. There are further possibilities for obvious or subtle stigmatization and structural disadvantages, such as access to research and progress for those with little equipment and skills in Big Data approaches. Algorithms can contain systematic biases, even ones which eventually contribute to racism. Having this in mind, it has to be assessed with due diligence in which areas and in which ways Big Data can contribute to discrimination. Adequate measures then have to be taken in order to prevent stigmatization and discrimination according to Article 11 of the UDBHR.

#### **IV.5. Sustainability with regard to energy and environment**

89. Big Data for health is also an ecological issue both for its energy cost and for the gains it can achieve through a better use of natural resources. Emphasis is put on the potential benefits, including energy savings and the fight against the greenhouse problem (Carbon War Room, n.d.). But energy-intensive information technologies are big consumers and produce greenhouse gases at all stages of their life cycle (CNRS, n.d.).

90. There are infrastructures behind a distributed and remote computing that have a high energy consumption and carbon footprint: 10% of world energy, health part unknown but most likely significant. This cost is comprised of 38% to terminals, 14% to transmission and 48% to data centers (Cappy, 2017). Of note, an estimated 80% of the energy consumption of connected objects (14 billion at present, 125 billion units in 2025) is not dedicated to their use but to maintaining their network connectivity. According to the International Energy Agency the world of connected devices consumed 616 TWh in 2013 (IEA, 2014), which is the consumption of Sweden with its 10 million inhabitants for one year. Data centers bring together thousands of machines running constantly. In addition to the energy required to operate the equipment, the huge needs for cooling must be added. In 2013, US data centers consumed an estimated 91 billion kilowatt-hours of electricity - which would be enough electricity to power all the households in New York City for two years - and are on-track to reach 140 billion kilowatt-hours by 2020 (NRDC, 2015).

91. Improvements might result from voltage equalization, research on quantum computing, machine architecture to solve the gap between data management and data storage, changing the voltage in microprocessors, as well as bio-inspiration which aims at improving the cost-effectiveness-ratio by learning from biological systems like the human brain.

92. There are also concerns about the possible impact on health by pollution associated with the waste created by the connected but no longer used objects. On a global scale, an estimated 50 million tons of E-waste are produced every year. The Stopping the E-waste Problem (StEP) initiative (<http://www.step-initiative.org/>), a joint effort from UN organizations, grassroots groups and industry, predicts that by 2017 the total annual volume of E-waste will have risen by a third, to 65.4 million tons.

93. There are also concerns with the ecological impact due to search and exploitation of essential components to produce around 67 million tons of new electrical and electronic equipment that are put on the market each year. The term 'rare earth' describes a set of 17 chemical elements with exceptional properties. It is thanks to these rare earths that colors of computer screens are so bright, mobile phones have touch screens and wind turbines can generate electricity. The downside is that their extraction and their transformation pollute, produce radioactive waste and distort the landscape. Indonesia and Australia have already been impacted.

#### **IV.6. Research**

94. Because Big Data can be a powerful generator of hypotheses, particularly for clinical trials, a primary ethical requisite is to conduct research following the highest standards of ethics, data protection and excellence. Biobanking is at the heart of biomedical research and forms an essential source for Big Data research. Accordingly, best practice guidance concerning Big Data research on biobanked samples and data collected and stored in academic biobanks have been already issued and are regularly discussed through recommendations of international organizations such as BBMRI-ERIC (<http://bbmri-eric.eu/>) and OECD. Recent advances include committees involving patient representatives reviewing the governance of the biobank including analysis of the use of the potential commercial value of the data, and a process of follow-up information on the various research using the data, allowing the patients to be informed long-term and giving them a real possibility to opt-out.

95. In an attempt to confirm the use of highest quality biobanked samples and associated data, the concept of BRIF (Bioresource Research Impact Factor) has been introduced. The BRIF initiative is intended to help in the follow-up of the use of data and databases associated with biological samples. In this context, the CoBRA (Citing of Bioresources in Research Articles) guideline has been developed to guide and standardise citation of bioresources in academic literature. This can be an example for Big Data approaches as well.

96. Understood in terms of 'control', ownership entails the right of data subjects to prevent any kind of data manipulation, including societal pressure to constrain any unacceptable uses (see Chapter IV.3.). For biobanks and databases, control is relevant, e.g. to consider the permissibility of using research data for commercial pursuits as is made possible by allowing private and third party companies to access biobanked samples and data. On the contrary, the ownership of the data may be an obstacle for research, for example preventing certain forms of data-mining or meta-analysis. National laws may limit such research, particularly if a transfer of data is involved.

97. The push towards open data access is in contradiction with protection of sensitive medically relevant data. In this context, it is noteworthy that some scientific journals require the transfer of all data sets used to obtain results to a repository in order to allow other scientists to replicate the results or to use them in meta-analysis. But this may also lead the publishers or other owners of those kind of repositories to take rights on the data sets and makes data-mining prohibited or too costly. Several national laws are presently under discussion to allow free access to data sets for research.

98. A major challenge is educating citizens to get the best from Big Data derived personal information. Firstly, subjects consenting to research must be able to exercise their access rights with reasonable effort so that the right can be exercised meaningfully. Secondly, Big Data requires advanced scientific knowledge through education in order to avoid

misunderstandings, e.g. with regard to the outcome of an analysis.

99. A risk exists of an overreliance in Big Data power which has led to the recent development of a direct relationship between data managers and patients/citizens bypassing medical intervention. Such practice does not exist in the academic field but is rapidly developing in private research by commercial companies. Medical interventions and interpretations are needed because predictions based on Big Data analysis alone do not cover all what needs to be covered. However, the interpretation of the meaning of Big Data results for a given person remains a fundamental goal of the relationship between patient and physician.

100. Big Data analyses will reveal results with potential health relevance for a given individual, which are sometimes referred to as incidental or unsolicited findings. Since these questions have been treated in a previous Report of the IBC on the Principle of Non-Discrimination and Non-Stigmatization (UNESCO, 2014), there is no need here to further comment.

101. At the level of public health research, Big Data will be used to develop public health policies to promote health and prevent disease at the collective level. Appropriate safeguards should be implemented to protect the rights to privacy of individuals and must adhere to good governance on the use of data. The cost of acquiring, maintaining and using Big Data in public health policy is also of major concern, especially in developing countries with scarce resources.

102. Health research now uses connected personal devices such as smartphones for tracking and monitoring to gather health data (mHealth). This can be a helpful source for real world data but it also raises ethical issues such as risk of exploitation or violation of the right to autonomy since mobile phones can collect a wide range and quantity of personal information from their user with few safeguards to protect autonomy and privacy. The use of such electronic devices for conducting health research and health surveys may still be efficient and effective as long as appropriate safeguards are implemented to protect the individual's autonomy and privacy.

## **V. GOVERNANCE**

103. Governance systems for Big Data should protect the fundamental rights of the persons from whom the data originates, including their freedom to make decisions, and aim at maintaining the trust of the public. Thus, guiding principles for governance have to include the respect of autonomy and the right to information, voluntariness, privacy and data protection, transparency, equality and lawfulness. Data governance should guarantee that citizen involvement, engagement, participation and sharing of data will not be subject to exploitation, manipulation and improper control. The IBC has based its recommendations on these principles.

104. Furthermore individual interests and rights as in public health policies can be balanced against what is called the public good. In the case of Big Data, most of the relevant actors collecting and using data generally lack a democratic mandate from the people whose information is included (willingly or unwillingly) in databases to define what the public good actually is about. This creates an ethical tension that must be dealt with by an appropriate governance structure which could guarantee a 'public good'.

105. In view of the global scope of Big Data and health, the huge variety of – ever new – players in this field and the fast technological development it is difficult to elaborate comprehensive and balanced regulations. But in order to make benefits possible, to ensure responsible conduct and to prevent harm prudent governance is required on different tiers and in different ways, guiding and binding the different stakeholders and institutions on the legal level as well as other levels in form of rules, regulations, decrees, agreements, self-binding code of conducts etc.

106. In the era of Big Data, it has become increasingly difficult to rely on the idea that data protection can be regulated solely by adhering to consent or anonymisation, without other safeguards (see also chapter on autonomy). This leads to the conclusion that it is crucial to develop and apply a comprehensive multi-tiered governance structure, which doesn't exist yet, for responsible use of data. It is also crucial to note, and has been noted already in paragraph 35, that while in this time and age the individual becomes more and more transparent because we know more and more about that individual, systems tend to become more and more opaque. This needs to be countered by creating more transparent governance systems. Such a governance system can follow different models, but in all cases should be a public structure, or at least a public-private partnership. Serious participation of patients and the public in setting up and applying the governance is a prerequisite for this.

107. Transparency about algorithms, which at the moment are mostly opaque, is an essential element in this context. Institutions or companies should be transparent about the algorithms that they use. With regard to algorithmic decision-making, they should offer explanations regarding both the procedures followed by the algorithm and the specific decisions that are made. Some even claim that the source code of an algorithm should be published. However, intellectual property rights as well as technological features, e.g. of self-learning algorithms, have to be taken into account when developing adequate schemes for transparency and control.

108. The fundamental right of data protection offers a conceptual and legal basis for guaranteeing a morally and legally acceptable use of data and for setting up a complex system of information governance rules. The principle of lawfulness, a key principle of data protection, requires that every processing of personal data needs a legitimate legal basis. Its essence is the overarching system of checks and balances, regulating personal data processing activities. It aims to complement the individual rights, guaranteed by the fundamental rights to privacy, with an effective allocation of responsibilities and duties of those who use the data. The principle of lawfulness has different legal status in different countries.

109. These governance framework should address regulatory and ethical issues related to every step of the cybernetic loop, including data collection, access, release, and linkage, analysis, reuse etc. Against the background of the ethical and legal considerations, the following elements are perceived to be crucial for such a framework.

- a. A clear statement of the purpose of the database.
- b. Procedures for (broad) consent, re-contact (including return of results) and re-consent.
- c. Procedures for dissent, as alternative to consent.
- d. Arrangements for ensuring the rights to access, to rectify, to cancel data.
- e. Arrangements for withdrawal, and a description of the extent to which withdrawal is technically possible.
- f. Arrangements for the protection of privacy, at least declaring the limits of privacy protection.
- g. Policies after the death of a participant.
- h. Arrangements on ownership of the data and the products derived from them.
- i. Arrangements for the collection, storage and duration of storage of data including quality control and safeguards to protect privacy and confidentiality.
- j. Arrangements for access to data including arrangements for data sharing, and criteria for access.
- k. Role of REC or DAC in decision-making about sharing of data.
- l. Arrangements for how the data will be dealt with in the event of change of ownership or closure of the database.

- m. Transparency of the algorithms used for pattern recognition; arrangements to check profiling of individuals or groups according to ethical considerations.
- n. Disclosure of commercial interest and collaboration with commercial parties.
- o. Arrangements that enable participants to remain informed of ongoing and novel use of data including research activities.
- p. A clear policy on the disclosure of aggregate and individual research results to participants.
- q. Arrangements for the involvement of participants in designing governance procedures, particularly with regard to ethical oversight and communication with data-providers.
- r. Arrangements for benefit-sharing.
- s. Provisions concerning indigenous / local populations and traditional minorities.
- t. Children and adolescents who reach the age of maturity during the research project should be given the opportunity to give informed consent for the continued storage and use of their data, and they should also be able to withdraw consent for future research.

110. These elements could be further developed into international treaties, self-binding regulations between providers or may even lead to the creation of a global agency of vigilance. Existing entities such as Research Ethics Committees (RECs), Data Access Committees (DACs) and data protection agencies could play an important role in such a multi-tiered governance framework, addressing relations between multiple stakeholders.

111. The IBC recommends that an International Legal Instrument on Data Protection is adopted by the Member States. If so requested, the IBC is willing to contribute to the development of a suitable governance framework that should precede such a treaty. Each country can then create within its own jurisdiction country specific legislation, under which an agency can be created that is tasked with the oversight of these governance systems. Such an agency would also provide a clear point of entry for public control.

## **VI. RECOMMENDATIONS**

112. In order to make Big Data a global success for health, to harvest the opportunities of Big Data in health care and research while at the same time avoiding violation of fundamental human rights enshrined in the Universal Declaration of Human Rights and in the Universal Declaration on Bioethics and Human Rights, the IBC gives the following recommendations.

113. Facing the complex nature, the global scope, and the wide variety of stakeholders involved in Big Data related to health, international cooperation as well as a multi-tiered governance approach are crucial, so that trust and control are balanced for the benefit of all (for more detail see Chapter V).

114. The IBC considers four measures to be crucial for protecting individual rights and fostering public good while recognizing the unavoidable loss of control by individuals about the use of their data in times of Big Data: governance, education, capacity building, and benefit sharing.

115. In the following, the IBC presents examples of important measures. They are understood as a step in an ongoing global debate. They address different stakeholders. The final success of these measures will depend on comprehensive cooperation and participation of all stakeholders including patients, research subjects, users and citizens in general.

116. International Agencies are called upon to develop and support a global framework for the use of Big Data in health related areas, especially in health care and research. For example:

- a. The UN is recommended to develop and adopt an International Legal Instrument on Data Protection in Health Care and Health Research.
- b. UNESCO is recommended to develop a convention on the protection of privacy, including a framework for new approaches to ownership and custodianship of personal data comprising but not only referring to health data. It should also address the handling of personal data and the digital presence of an individual after his or her physical death. This convention can build on the draft resolution of the Human Rights Council on the right to privacy in the digital age (A/HRC/34/L.7/Rev.1).
- c. WHO is recommended to encourage the establishment of an agreement by app-stores about presenting health-related apps in a way that autonomy, transparency and adequate information are guaranteed.
- d. A global vigilance system for the use of Big Data in health related applications should be set up.
- e. International Agreements should adopt an understanding of Big Data as a common good of humankind and facilitate open access and use of Big Data for the common good where feasible. For this purpose a public data infrastructure should be set up.
- f. OECD is recommended to develop a framework for sharing of benefits from Big Data applications. This should also contribute to increasingly overcome the digital gap.
- g. International institutions as well as regional and national institutions that lay down technical standards for devices and processes in the context of Big Data are called upon to define these standards according to the ethical principles of respect for autonomy, privacy, and justice as elaborated in this report, as well as to the need for transparency.
- h. The International Energy Agency (IEA) is called upon to synergize efforts for a sustainable and responsible use of energy in Big Data management. It could set up programs to develop energy efficiency policies including dialogue with non-member countries.
- i. Environmental Protection Agencies such as UN Environment together with the WHO are called upon to set up a coordinated action plan for saving rare resources. Additionally they should implement a program to avoid E-waste impacting the health of people especially in developing countries.

117. National Governments are called upon to develop and launch an action plan including legislation and policies which should among others address the following aspects:

- a. Internationally harmonized implementation of globally accepted data protection principles.
- b. National governments – in so far as not done already - are called upon to establish an effective Data Controlling Agency. All national Data Controlling Agencies should work together with the above mentioned (paragraph 116[d]) Global Vigilance System for the use of Big Data in health related applications. This coordinated system could serve as a starting point for the International Legal Instrument on Data Protection in Health Care and Health Research.
- c. Capacity-building for Big Data-related health care and research including an efficient data infrastructure.
- d. Promotion of learning health care systems which use data from everyday health care to produce evidence for future health care.
- e. Enabling effective cross-border cooperation in the processing of personal health data for health-related public interest purposes.

- f. Promotion of education with regard to Big Data-related skills, competences, and awareness about ethically relevant implications. Particular attention should be paid to vulnerable groups like minors and people with impaired capacities.
- g. Introduction of diverse models of consent which are specified according to the context of health care and research, allowing for broad and dynamic consent where appropriate, feasible and safeguarded. The same applies to electronic consent.
- h. Protection of privacy of individuals sharing their data by default as well as by design of the device and technological equipment. Separate attention should be paid to the protection of group privacy – not least against the background of the manifold possibilities of discrimination.
- i. Implementation of a Research Ethics Committee or a similar institution overseeing commercial Big Data research.
- j. Development and implementation of instruments for meaningful public and patient involvement in a culture sensitive way.
- k. Governments should launch coordinated programs that motivate device developers to create communication protocols that enable energy savings. A priority should be to standardize procedures for data collection and management.

118. The wide variety of health care and research institutions as well as companies, are called upon to develop context sensitive guidelines, code of conducts and self-binding instruments with a special focus on Big Data applications. Ethical policies and codes of conduct should also be developed for health professionals, computer scientists, clinical researchers, data scientists and other stakeholders. Motivating ethical behavior should be the primary goal, but consequences should also be provided for in case ethical principles are violated.



## BIBLIOGRAPHY

African Union (AU). 2014. *African Union Convention on Cybersecurity and Personal Data Protection*. Addis Ababa, AU. Available at: <https://www.au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection>

Andanda, P., Wathuta, J., Leisinger, K. and Schroeder, D. 2016. *National and International Compliance Tools*, a report for the TRUST Project. Available at: <http://trust-project.eu/wp-content/uploads/2017/02/TRUST-664771-National-and-International-Compliance-Tools-Final.pdf>

Asian-Pacific Economic Cooperation (APEC). 2015. *Updates to the APEC Privacy Framework*. Singapore, APEC Secretariat. Available at: [http://mddb.apec.org/Documents/2016/SOM/CSOM/16\\_csom\\_012app17.pdf](http://mddb.apec.org/Documents/2016/SOM/CSOM/16_csom_012app17.pdf)

Bowker, G. C. 2014. Big data, big questions the theory/data thing, *International Journal of Communication*. Los Angeles, USC Annenberg Press, Vol. 8, pp. 1795-1799.

Cappy, A. 2017. Impact énergétique des Big Data (Energy Impact of Big Data). In: Bouzeghoub, M. and Mosseri, R. eds. *Les Big Data à découvert* (Big Data Uncovered). Paris, CNRS.

Carbon War Room. n.d. *Carbon War Room* (online). Available at: <http://carbonwarroom.com/>

Centre national de la recherche scientifique (CNRS). n.d. *EcolInfo* (online). Paris, CNRS. Available at: <http://ecoinfo.cnrs.fr/>

Choudhury, S., Fishman, J. R., McGowan, M. L., and Juengst, E. T. 2014. Big Data, open science and the brain: lessons learned from genomics, *Frontiers in Human Neuroscience*. Lausanne, Frontiers, 8, 239. doi:10.3389/fnhum.2014.00239

Cohen, J. E. 2013. What privacy is for, *Harvard Law Rev.* Cambridge, Harvard Law Review, Vol. 126, No. 7, pp. 1904-1933.

Coles, D., Wathuta, J. and Andanda, P. 2016. ICT and Mobile Data for Health Research. In: Schroeder, D., Cook Lucas, J., Fenet, S. and Hirsch, F. eds. *"Ethics Dumping" – Paradigmatic Case Studies*, a report for the TRUST Project. Available at: <http://trust-project.eu/wp-content/uploads/2016/03/TRUST-664771-Paradigmatic-Case-Studies-WP1-Final.pdf>

Comitato Nazionale per la Bioetica (CNB). 2016. *Information and Communication Technologies and Big Data: Bioethical Issues*. Rome, CNB. Available at: [http://bioetica.governo.it/media/172149/p124\\_2016\\_information-technologies-and-big-data\\_en.pdf](http://bioetica.governo.it/media/172149/p124_2016_information-technologies-and-big-data_en.pdf)

Coordinated Research Infrastructures Building Enduring Life-science Services (CORBEL). n.d. Available at: <http://www.corbel-project.eu/home.html>

Council of Europe (CoE). 1950. *Convention for the Protection of Human Rights and Fundamental Freedoms* (European Convention on Human Rights). Strasbourg, CoE. Available at: <http://www.echr.coe.int/pages/home.aspx?p=basictexts>

CoE. 1981. *Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data*. Strasbourg, CoE. Available at: <http://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/108>

CoE. 2001. *Additional Protocol to the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, regarding supervisory authorities and transborder data flows*. Strasbourg, CoE. Available at: <http://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/181>

European Commission (EC). 2012. *Opinion No. 26: Ethics of Information and Communication Technologies*, European Group on Ethics in Science and New Technologies (EGE). Luxembourg, Publications Office of the European Union. Available at: <http://bookshop.europa.eu/en/ethics-of-information-and-communication-technologies-pbNJAJ12026/>

European Commission (EC). 2015. *Opinion No. 29: The ethical implications of new health technologies and citizen participation*, European Group on Ethics in Science and New Technologies (EGE). Luxembourg, Publications Office of the European Union. Available at: <https://publications.europa.eu/en/publication-detail/-/publication/e86c21fa-ef2f-11e5-8529-01aa75ed71a1/language-en/format-PDF/source-39541144>

European Court of Justice (ECJ). 2012. *Judgement of the Court (Grand Chamber): Used Soft GmbH v Oracle International Corp (C-128/11)*. Luxembourg, ECJ. Available at: <http://curia.europa.eu/juris/document/document.jsf?docid=124564&doclang=EN>

European Medicines Agency (EMA). 2016. *External guidance on the implementation of the European Medicines Agency policy on the publication of clinical data for medicinal products for human use*. London, EMA. Available at: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Regulatory\\_and\\_procedural\\_guideline/2016/03/WC500202621.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2016/03/WC500202621.pdf)

European Union (EU). 1995. *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data*. Official Journal of the European Union (online). Available at: <http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:31995L0046>

EU. 1996. *Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases*. Official Journal of the European Union (online). Available at: <http://eur-lex.europa.eu/legal-content/FR/TXT/?uri=celex%3A31996L0009>

EU. 2015. *General Data Protection Regulation (compromise text after trilogue)* (online). Available at: <http://www.statewatch.org/news/2015/dec/eu-council-dp-reg-draft-final-compromise-15039-15.pdf>

EU. 2016. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. Official Journal of the European Union (online). Available at: <http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32016R0679>

H3Africa Consortium. 2014. Research capacity. Enabling the genomic revolution in Africa, *Science*. New York, American Association for the Advancement of Science (AAAS), Vol. 344, No. 6190, pp. 1346-1348.

Hoeren, T. 2014. Big Data and the Ownership in Data: Recent Developments in Europe, *European Intellectual Property Review*. London, Sweet & Maxwell, Vol. 36, No. 12, pp. 751-754.

Innovative Medicines Initiative (IMI). 2014. *Code of Practice on Secondary Use of Medical Data in Scientific Research Projects*, final draft 27 August 2014. Available at: [http://www.imi.europa.eu/sites/default/files/uploads/documents/CodeofPractice\\_SecondaryUseDRAFT.pdf](http://www.imi.europa.eu/sites/default/files/uploads/documents/CodeofPractice_SecondaryUseDRAFT.pdf)

Institute of Medicine (IOM). 2007. *The Learning Healthcare System: Workshop Summary*. Washington D.C., The National Academies Press.

IOM. 2013. *Best care at lower cost: The path to continuously learning health care in America*. Washington D.C., The National Academies Press.

International Data Corporation (IDC). 2014. *The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things, Executive Summary: Data Growth, Business Opportunities, and the IT Imperatives* (online). Available at: <https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm>

International Energy Agency (IEA). 2014. *More Data, Less Energy: Making Network Standby More Efficient in Billions of Connected Devices*. Paris, IEA. Available at: [https://www.iea.org/publications/freepublications/publication/MoreData\\_LessEnergy.pdf](https://www.iea.org/publications/freepublications/publication/MoreData_LessEnergy.pdf)

International Telecommunication Union (ITU). 2015a. *ICT Facts and Figures 2015*. Geneva, ITU. Available at: <http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2015.pdf>

ITU. 2015b. *ITU's Annual Report 2015*. Geneva, ITU. Available at: <http://www.itu.int/en/annual-report-2015/Pages/default.aspx>

ITU. 2016. *ITU's Annual Report 2016*. Geneva, ITU. Available at: <http://www.itu.int/en/annual-report-2016/Pages/default.aspx>

ITU. 2016. *ICT Facts and Figures 2016*. Geneva, ITU. Available at: <http://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx>

ITU. 2017. *ICT Facts and Figures 2017*. Geneva, ITU. Available at: <http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf>

Kaplan, W., Wirtz, V. J., Mantel-Teeuwisse, A., Stokl, P., Duthey, B. and Laing, R. 2013. *Priority Medicines for Europe and the World: 2013 Update*. Geneva, WHO. Available at: [http://www.who.int/medicines/areas/priority\\_medicines/MasterDocJune28\\_FINAL\\_Web.pdf?ua=1](http://www.who.int/medicines/areas/priority_medicines/MasterDocJune28_FINAL_Web.pdf?ua=1)

Lavenex S. and Schimmelfennig F. 2009. EU rules beyond EU borders: theorizing external governance in European politics, *Journal of European public policy*. Taylor and Francis Online, Vol. 16, No. 6, pp. 791-812.

Laurie, G., Ainsworth, J., Cunningham, J., Dobbs, C., Jones, K. H., Kalra, D., Lea, N. C. and Sethi, N. 2015. On moving targets and magic bullets: Can the UK lead the way with responsible data linkage for health research?, *International Journal of Medical Informatics*. Amsterdam, Elsevier, Vol. 84, pp. 933-940.

Lea, N. C. 2015. *Design and Development of a Knowledge Modelling Approach to Govern the Use of Electronic Health Records for Research*. Doctoral thesis, University of London, UK.

Mertz M., Jannes, M., Schlomann, A., Manderscheid, E., Rietz, C. and Woopen, C. 2016. *Digitale Selbstbestimmung*. Cologne Center for Ethics, Rights, Economics, and Social Sciences of Health. DOI 10.8716/ceres/00001,

Mittelstadt, B. D. and Floridi L. 2016. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts, *Sci Eng Ethics*. New York, Springer, Vol. 22, pp. 303-341.

Mostert, M., Bredenoord, A. L., Biesart, M. C. I. H. and van Delden, J. J. M. 2015. Big Data in medical research and EU data protection law: challenges to the consent or anonymise approach, *European Journal of Human Genetics*. London, Nature Publishing Group, Vol. 24, No. 7, pp. 956-960. doi:10.1038/ejhg.2015.239

Natural Resources Defense Council (NRDC). 2015. *America's Data Centers Consuming and Wasting Growing Amounts of Energy* (online). New York, NRDC. Available at: <https://www.nrdc.org/resources/americas-data-centers-consuming-and-wasting-growing-amounts-energy>

Nielsen, M. 2014. *Who Owns Big Data?* (online). BBVA OpenMind. Available at: <https://www.bbvaopenmind.com/wp-content/uploads/2014/02/BBVA-OpenMind-Technology-Innovation-Internet-Bussines-Michael-Nielsen-Who-Owns-Biq-Data.pdf>

Nuffield Council on Bioethics. 2015. *The collection, linking and use of data in biomedical research and health care: ethical issues*. London, Nuffield Council on Bioethics. Available at: [http://nuffieldbioethics.org/wp-content/uploads/Biological\\_and\\_health\\_data\\_web.pdf](http://nuffieldbioethics.org/wp-content/uploads/Biological_and_health_data_web.pdf)

Organisation for Economic Cooperation and Development (OECD). 2010. *Improving Health Sector Efficiency: The Role of Information and Communication Technologies*. Paris, OECD. Available at: <http://www.oecd.org/els/health-systems/improvinghealthsectorefficiency.htm>

OECD. 2013a. *Guidelines Governing the Protection of Privacy and Transborder Flow of Personal Data*. Paris, OECD. Available at: <https://www.oecd.org/sti/ieconomy/2013-oecd-privacy-guidelines.pdf>

OECD. 2013b. *Strengthening Health Information Infrastructure for Health Care Quality Governance: Good Practices, New Opportunities and Data Protection Challenges*. Paris, OECD. Available at: <http://www.oecd.org/publications/strengthening-health-information-infrastructure-for-health-care-quality-governance-9789264193505-en.htm>

OECD. 2015. *Health Data Governance: Privacy, Monitoring and Research*. Paris, OECD. Available at: <http://www.oecd.org/fr/publications/health-data-governance-9789264244566-en.htm>

OECD. 2017. *OECD Recommendation on Health Data Governance*. Paris, OECD. Available at: <http://www.oecd.org/els/health-systems/health-data-governance.htm>

Pharmaceutical Research and Manufacturers of America (PhRMA)/European Federation of Pharmaceutical Industries and Associations (efpia). 2013. *Principles for Responsible Clinical Trial Data Sharing* (online). Available at: <http://transparency.efpia.eu/uploads/Modules/Documents/data-sharing-prin-final.pdf>

Rathenau Instituut. 2017. *Human rights in the robot age. Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality*. The Hague, Rathenau Instituut. Available at: <https://www.rathenau.nl/en/publication/human-rights-robot-age-challenges-arising-use-robotics-artificial-intelligence-and>.

Science International. n.d. *Open Data in a Big Data World*. Paris, International Council for Science (ICSU), International Social Science Council (ISSC), the World Academy of Sciences (TWAS) and the InterAcademy Partnership (IAP). Available at: <http://www.science-international.org/>

Sethi, N. and Laurie, G. 2013. Delivering proportionate governance in the era of eHealth: making linkage and privacy work together, *Med Law Int*. Thousand Oaks, SAGE Publications, Vol. 13, No. 2-3, pp. 168–204.

Shilton, K. 2009. Four billion little brothers? Privacy, mobile phones, and ubiquitous data collection, *Communications of the ACM*. New York, ACM, Vol. 52, No. 11, pp. 48-53.

United Nations (UN). 1948. *Universal Declaration of Human Rights*. New York, UN. Available at: <http://www.un.org/en/universal-declaration-human-rights/index.html>

UN. 1990. *Guidelines for the Regulation of Computerized Personal Data Files*, adopted by the General Assembly resolution 45/95 of 14 December 1990. New York, UN. Available at: <http://www.refworld.org/pdfid/3ddcafaac.pdf>

United Nations Conference on Trade and Development (UNCTAD). 2016. *Data protection regulations and international data flows: Implications for trade and development*. Geneva, UNCTAD. Available at: [http://unctad.org/en/PublicationsLibrary/dtlstict2016d1\\_en.pdf](http://unctad.org/en/PublicationsLibrary/dtlstict2016d1_en.pdf)

United Nations Educational, Scientific and Cultural Organization (UNESCO). 2005. *Universal Declaration on Bioethics and Human Rights*. Paris, UNESCO. Available at: [http://portal.unesco.org/en/ev.php-URL\\_ID=31058&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/en/ev.php-URL_ID=31058&URL_DO=DO_TOPIC&URL_SECTION=201.html)

UNESCO. 2014. *Report of the International Bioethics Committee of UNESCO (IBC) on the Principle of Non-discrimination and Non-stigmatization*. Paris, UNESCO. Available at: <http://unesdoc.unesco.org/images/0022/002211/221196E.pdf>

UNESCO. 2015. *Report of the International Bioethics Committee of UNESCO (IBC) on the Principle of the Sharing of Benefits*. Paris, UNESCO. Available at: <http://unesdoc.unesco.org/images/0023/002332/233230E.pdf>

United States of America (USA). 1996. *Health Insurance Portability and Accountability Act*. Washington D.C., US Congress. Available at: <https://www.gpo.gov/fdsys/pkg/PLAW-104publ191/pdf/PLAW-104publ191.pdf>

USA. 2002. *HIPAA Privacy Rule*. Washington D.C., US Department of Health and Human Services. Available at: <https://www.hhs.gov/hipaa/for-professionals/privacy/index.html?language=es>

USA. 2009. *The American Recovery and Reinvestment Act*. Washington D.C., US Congress. Available at: <https://www.gpo.gov/fdsys/pkg/BILLS-111hr1enr/pdf/BILLS-111hr1enr.pdf>

USA. 2014. *Big Data: Seizing Opportunities, Preserving Values*. Washington D.C., Executive Office of the President. Available at: [https://obamawhitehouse.archives.gov/sites/default/files/docs/big\\_data\\_privacy\\_report\\_5.1.1\\_4\\_final\\_print.pdf](https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_5.1.1_4_final_print.pdf)

USA. n.d. *Privacy Shield Program Overview* (online). Washington D.C., US Department of Commerce. Available at: <https://www.privacyshield.gov/Program-Overview>

Vayena, E. and Tasioulas, J. 2015. "We the Scientists": a Human Right to Citizen Science. *Philos. Technol.* New York, Springer, Vol. 28, pp. 479-485

Weiss, M. A. and Archick, K. 2016. *U.S.-EU Data Privacy: From Safe Harbor to Privacy Shield*. Washington D.C., US Congressional Research Service. Available at: <https://fas.org/sqp/crs/misc/R44257.pdf>

World Health Organization (WHO). 2011. *mHealth: New horizons for health through mobile technologies: second global survey on eHealth*, Global Observatory for eHealth series-Volume 3. Geneva, WHO. Available at: [http://apps.who.int/iris/bitstream/10665/44607/1/9789241564250\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/44607/1/9789241564250_eng.pdf)

WHO. n.d. *eHealth* (online). Available at: <http://www.who.int/ehealth/en/>

World Medical Association (WMA). 2013. *WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects* (online). Available at: <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>

WMA. 2016. *WMA Declaration of Taipei on Ethical Considerations Regarding Health Databases and Biobanks* (online). Available at: <https://www.wma.net/policies-post/wma-declaration-of-taipei-on-ethical-considerations-regarding-health-databases-and-biobanks/>

World Trade Organization (WTO). 1994. *Agreement on Trade-Related Aspects of Intellectual Property Rights* (TRIPS). Geneva, WTO. Available at: [https://www.wto.org/english/tratop\\_e/trips\\_e/intel2\\_e.htm](https://www.wto.org/english/tratop_e/trips_e/intel2_e.htm)

Wyber, R., Vaillancourt, S., Perry, W., Mannava, P., Folaranmi, T. and Celi, L.A. 2015. Big data in global health: improving health in low- and middle-income countries, *Bulletin of the World Health Organization*. Geneva, WHO, Vol. 93, pp. 203-208. doi: <http://dx.doi.org/10.2471/BLT.14.139022>