UNIVERSITY OF SOUTHAMPTON

Faculty of Physical Sciences and Engineering

School of Electronics and Computer Science

What Amendments Need to Be Made to the Current EU Legal Framework to Better Address the Security Obligations of Data Controllers?

by Evangelia Papadaki

Supervisors:
Dr Kieron O'Hara (Faculty of Physical Sciences and Engineering)
Dr Sophie Stalla-Bourdillon (Faculty of Business, Law and Art)

A thesis submitted in partial fulfilment for the
degree of Doctor of Philosophy

March 2018

Abstract

The overall objective of this thesis is to identify the gaps in the current EU legal framework surrounding the security obligations of data controllers and make recommendations to help advance the discussions around the possible ways of effectively addressing the problem of cyber insecurity. The thesis adopts an interdisciplinary approach to data security, which involves legal analysis enriched with considerations from the fields of Computer Science and Managerial Economics. In response to the rapidly changing landscape of emerging technologies, which challenges the conventional thinking of regulators, the thesis calls for a shift in the data security regulation paradigm. The contribution of the thesis to knowledge in this field lies in reframing the elements that need to be incorporated into the laws regulating the security obligations of data controllers. The thesis proposes a holistic, dynamic, hybrid and layered approach to data security, which systematically tailors the security obligations of data controllers to the level of re-identification risk involved in data processing operations, and suggests security measures depending on the security level required while laying down the security objectives to be achieved. The proposed regulatory model can serve as guidance for regulators on the law-making process concerning the security obligations of data controllers. The proposed model aspires to provide adequate clarity to data controllers in terms of the initial phase of the design of security measures, while abstaining from imposing technology specific security requirements in order to grant flexibility to data controllers to adapt the security mechanisms to their particular business model and the given data environment.

Contents

Academic Thesis: Declaration of Authorship

I, Evangelia Papadaki, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

"What Amendments Need to Be Made to The Current EU Legal Framework to Better Address The Security Obligations of Data Controllers?"

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Either none of this work has been published before submission, or parts of this work have been published as: [please list references below]:

Signed: EVANGELIA PAPADAKI

Date: 27/03/2018

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr Sophie Stalla-Bourdillon for her continuous support, immense knowledge, and our stimulating discussions. I would also like to thank the rest of my PhD supervisory team, Dr Kieron O'Hara and Dr Tim Chown, for their insightful comments and expertise that greatly assisted my research. This research was supported by the EPSRC Centre for Doctoral Training and the Web Science Institute at the University of Southampton and I am grateful to the people who provided me the opportunity to explore in depth such an exciting field of research. Last but not least, I would like to thank my family and friends for supporting spiritually this challenging and fascinating PhD journey.

CHAPTER 1: Introduction

Nowadays cyber insecurity is considered to be a "monumental problem" as security breaches may severely impact not only governments, companies and individuals, but may also generate substantial financial losses for the wider economy and negatively affect societal welfare (European Commission, 2013). Although there is no common understanding of the term 'cyber security' and, therefore, no general consensus on a common definition, since different definitions have been suggested by different actors,[1] the essence of cyber security lies in putting in place procedures and measures able to protect both the physical elements (such as physical infrastructure and buildings) and the virtual elements (such as networks and data) of the cyber environment. In the context of an organisation's cyber environment, cyber security ensures the attainment and maintenance of the security properties of the organisation's assets (telecommunications systems, connected computing devices, applications, services, stored and transmitted information) against relevant risks presented in the cyber environment (Lie et al., 2009; UK Government, 2013). The fact that a set of Internet problems, such as hacking, phishing, and viruses, collectively can be characterised as security harms that first appeared with the emergence of the Internet since the conduct involved is motivated by the rise of the heavily interconnected networks that comprise the Internet, has resulted in network security to be considered as the most important aspect of cyber security. Network security refers to the process of implementing physical and software preventative measures to protect the underlying networking infrastructure from unauthorised access, misuse, modification, or destruction. The aim of network security is to assure that the network performs its critical functions correctly and to preserve the confidentiality (ensuring the protection of data from unauthorised access), integrity (ensuring that data are protected against unauthorised modification and/or destruction) and availability (ensuring that data are accessible when needed) of all systems and information on the network (Liska, 2003; Harrington, 2005; Douligeris & Serpanos, 2006; Alpcan & Basar, 2011; Stewart, 2014). As indicated, the term 'network security' is closely related to the term 'information security', which is also evident in the definition provided in Recital 49 of the General Data Protection Regulation (hereinafter referred to as 'GDPR'),[2] which describes network and information

[1] *See* Deepak Rout, 2015, Developing a Common Understanding of Cyber security, *ISACA Journal* Vol. 6, available at:https://www.isaca.org/Journal/archives/2015/volume-6/Pages/developing-a-common-understanding-of-cybersecurity.aspx.

[2] 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

security as "the ability of a network […] to resist, at a given level of confidence, accidental events or unlawful or malicious actions that compromise the availability, authenticity, integrity and confidentiality of stored or transmitted data, and the security of the related services offered by, or accessible via, these networks […], by public authorities, Computer Emergency Response Teams (CERTs), Computer Security Incident Response Teams (CSIRTs), providers of electronic communications networks and services and services, and by providers of security technologies and services". The main elements of the definition of network security, that is, the availability, authenticity, integrity and confidentiality of data, constitute the core of the term 'information security', which should be understood as the necessary precautions that must be applied to data, whether it is its creation, its use, its backup, its archiving or destruction, in order to prevent its alteration and damage, or access by non-authorised third parties (CNIL, 2010; ENISA, 2014).

Given that the majority of current cyber-attacks share the common goal of acquiring confidential information - the gold of the twentieth-first century – and, as a consequence, are directed at companies whose servers store vast amounts of data, such as credit card and personal information of their customers, the role of organisations in charge of the processing of personal data in improving the level of cyber security proves to be of crucial significance. Recent years have seen an exponential increase in the generation, collection, analysis and exchange of personal information stemming from the emergence of technological innovations such as the Internet of things, cloud computing, Open data and Big data. The fact that companies handling personal data are the primary target of cyber criminals, coupled with their ability of developing highly detailed profiles of their customers provided by the aforementioned technologies or even technologies used for security purposes, such as the case of deep packet inspection, implies the existence of some sort of duty of care of such organisations best suited to protect their customers' personal data. However, the increasing number of widespread security incidents of unauthorised or improper use and sharing of personal information signals the failure of companies to put in place the necessary security measures that would have prevented such incidents from occurring due to a number of barriers acting against widespread adoption of effective cyber risk management by organisations. Organisations often lack sufficiently clear standards according to which to operate or may not have adequate economic or legal incentives to curtail security breaches

repealing Directive 95/46/EC (General Data Protection Regulation). The GDPR will take effect on 25 May 2018 replacing the current Data Protection Directive and will be directly applicable in all Member States without the need for implementing national legislation.

(FTC, 2012). Although it is expected that organisations would be motivated to protect their own sensitive data and online presence, it may not be in their commercial interests to mitigate against the wider external costs that could occur from a successful attack that affects personal information (i.e. of customers or employees), or other businesses' commercial information, or maybe organisations do not fully comprehend the potential negative consequences of a security breach (UK Government, 2016). The combination of the above deterring factors is likely to lead to organisations under-investing in a sufficient level of cyber protection. In this regard, there is a strong justification for government intervention on the grounds of a clear public interest in protecting citizens' personal data, contrary to the commercial interests of companies that do not necessarily involve implementing protection to a level that is in the public interest. Therefore, government has a clear role to play in addressing the aforementioned barriers and ensuring that market incentives work to maximise data security. While data security is a shared responsibility of government, the private sector and individuals, only when government acts as a catalyst that bridges the interests' gap between the organisations and individuals, can data security be comprehensively addressed. In particular, the regulator's role in data security consists in identifying and evaluating potential risks and threats relating to personal data, establishing data security objectives and setting out the roles and responsibilities of all stakeholders in the process. It is essential that data security regulation provides legal certainty and hence serve as a key mechanism for building trust between individuals and the organisations handling their personal data.

In order to help advance the discussions around the possible ways of effectively addressing the problem of cyber insecurity, this thesis explores the role of data protection regulation in the EU context. In particular, in an attempt to explain the paradox of constantly increasing data breaches despite the existence of data protection laws, the thesis exposes limitations and gaps in the EU legal framework surrounding the security obligations of data controllers.[3] At this point, it should be noted that the role of various other parameters responsible for the fact that cyber insecurity is still an open problem has not been neglected, but these factors are outside the scope of this thesis, which examines the cyber insecurity problem from a legal perspective while enriching the legal analysis with considerations from the fields of Computer Science and Managerial Economics. Another point to be made is that the analysis

[3] *'Data controller'*, as defined in the EU data protection legal framework, is "the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data" (Article 2(d) of the Data Protection Directive).

is primarily based on the Data Protection Directive (hereinafter referred to as 'DPD')[4], which serves as a point of reference for the other pieces of legislation examined. Moreover, the thesis excludes from its scope the NIS Directive[5] as the security requirements stipulated do not specifically deal with the processing of personal data.

The main research question the thesis aims to answer is what amendments need to be made to the current EU legal framework to better address the security obligations of data controllers. To this end, the thesis is divided in two sections, where the first section deals with the analysis of the current data security regulation paradigm, while the second section explains the reasons due to which the current paradigm has failed to address the problem and suggests solutions. More specifically, the first part of the thesis revolves around the concept of technological neutrality aiming to unpack its meaning by studying the reasons that led the EU legislator to apply this concept to the data security context. The emergence of information and communications technologies (ICTs) has prompted the demand for new regulatory regimes to adopt a level of abstraction and thus abstain from targeting specific technologies to be regulated or used for the implementation of certain provisions. Technological neutrality has long been held up as a guiding principle for the regulation of technology in the field of ICT regulation and continues to be a pervasive concept in this field. Since 2011, technological neutrality has also been recognised as a key principle for Internet policy (OECD, 2011). Although the principle of technology neutrality is widely accepted as it often figures in policy documents and regulatory instruments, its meaning is not always clearly understood. The legal uncertainty, in which technology neutral approaches may result, as well as the discretion granted to data controllers in terms of their security obligations, are examined as possible culprits for the reluctance on the part of companies to sufficiently safeguard individuals' personal data. In addition, the analysis of the hazards hidden in a blind adoption of the technology neutral approach is indicative of the need for regulators to weigh up the pros and cons of this principle. The second part unfolds the contribution of the thesis to the existing work in this field, which lies in the fact that it calls for a paradigm shift in the way data security regulation is developed reframing the elements that need to be incorporated into the laws regulating the security obligations of data controllers, and proposing an

[4] 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

[5] Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union

alternative regulatory model. The thesis aims to address the problem of the lack of incentives for the companies to adopt robust security mechanisms in the current legal framework by building a regulatory model able to align companies' interests with those of individuals incorporating industry's practices into the law-making process, which also results in laws more easily adaptable to technological changes. Furthermore, the thesis strengthens the existing literature by proposing a hybrid and layered approach to data security, which systematically tailors the security obligations of data controllers to the level of re-identification risk for different states of data and suggests security measures depending on the security level required. Finally, the proposed model, which can serve as guidance to direct lawmakers towards a holistic approach to cyber security, lays down security objectives, that is, security outcomes that should be achieved, which cover all the aspects of cyber security.

Chapter 2 articulates the research approach (epistemology and norms) undertaken for this style of research and how it was implemented for the work presented in this thesis. Chapter 3 analyses the legal approach to data security adopted at the EU level and identifies any divergences between the jurisdictions of Member States. In particular, this chapter examines whether the EU data protection laws specify any types of security obligations imposed on data controllers to answer the question of whether it is enough for data controllers to implement the security obligations incumbent on them or there is a need for defining more detailed obligations. The widely adopted technology approach to data protection is thoroughly analysed in the context of the EU data protection regime and the comparison between Member States' national laws reveal the divergent approaches adopted in the implementation of the DPD. To identify any possible amendments brought about to the security requirements of data controllers by the reform of the data protection legal framework, the GDPR is studied and the extent to which the concept of 'data protection by design' included in the Regulation violates the principle of technology neutrality is examined. The following three chapters deal with the research question of which legislative approach would be more suitable in the data security context, while also trying to identify the hazards and weaknesses of the conceptual foundations on which the current legal framework has been established. Chapter 4 provides insights on the mindset underlying the EU data protection legal framework and answers the question of whether the DPD meets the conditions required for a legal approach to be considered as technology neutral. The elements of the definitions of 'personal data' and 'data processing' reveal the technological assumptions hidden in the approach adopted by the EU legislator and thus undermine the seemingly technology neutral

character of the DPD. Chapter 5 examines the technology neutral approach with respect to the EU legal framework regulating electronic signatures to reveal the technology favourism that might be hidden in technology neutral laws, which, in turn, demonstrates the difficulties in trying to adopt a regulatory approach able to accommodate different kinds of future technologies. The concerns stemming from technology neutral laws in the field of surveillance and data processing activities, where technical considerations with severe practical implications are not included in primary legislation but are rather left to secondary legislation, are indicative of the excessive powers granted in case technological neutrality leads to technological ambiguity. Chapter 6 also focusses on the data controllers' discretion when applying security measures and, more specifically, the technique of deep packet inspection, whose application might severely jeopardise individuals' privacy if used inappropriately. After presenting examples indicative of the problematic nature of deep packet inspection, the chapter proceeds with the legal safeguards that should be taken into consideration when making laws concerning techniques that have the potential to impact on individuals' fundamental rights and violate the basic principles of Internet governance. The last two chapters tackle the following questions: whether the reluctance of the private sector to implement robust security mechanisms could be attributed to the current legal approach to data security (or else, whether the private sector is provided with sufficient motives to abide by the legal security requirements); how to overcome the deficiencies of the current legal framework; how to change the philosophy underlying the technology-related laws; how should the law-making process be reformed. Chapter 7 examines the feasibility of the development of a regulatory system able to align the private sector's interests with those of individuals while abstaining from impeding technological innovation. The role of performance standards is analysed and suggestions are made as to the integration of industry's practices, such as the adaptive management process, into the law-making process as potential means of striking the right balance between two seemingly irreconcilable concepts – regulation and innovation. The concept of smart regulation is also examined and emphasis is placed on the need for its integration through the entire law-making cycle. Chapter 8 aims to build a regulatory model able to address the issues arising from the problematic nature of the identifiability concept in the current EU legal framework surrounding the security obligations of data controllers. The first part of the chapter examines the benefits of the risk-based approach to data protection and the notion of risk as it appears in the provisions of the GDPR, according to which data processing operations which involve high risk trigger additional compliance obligations. The second part provides insights on the

need for developing a regulatory model able to overcome the challenges posed by the binary distinction between personal and anonymous data in the post-anonymisation era, which will provide the necessary incentives for companies to implement adequate security mechanisms. Drawing on the literature on the degrees of data identifiability, a regulatory model is proposed that aims to provide clarity to data controllers' security obligations by identifying the most crucial risk vectors, laying down high-level security objectives and recommending examples of security measures to be implemented according to the security level required in each case.

CHAPTER 2: Methodology

2.1. Introduction to the Research Approach Undertaken

The overall goals of this thesis are firstly to establish the significance of the general field of study and then identify a place where a new contribution could be made. In other words, the aim is to identify the knowledge gap by moving from the general/known to the specific/unknown. To fill the possible knowledge gaps, and thus contribute to the overall knowledge on the research topic, literature review has served as a nexus to link the new information to the known body of knowledge by giving insights to what has been written about the issue at stake, what are the missing aspects and possible gaps in the literature. Arguments have been derived from both primary sources, such as existing laws, and secondary sources including scholarly publications, government reports, parliamentary committees' reports, policy documents, international guidelines, law reform documents, journal articles, media reports, and technical standards. In order to examine the problem of cyber security, and more specifically, the reasons why the number of cyber-attacks is constantly increasing despite the existence of laws regulating data security requirements, the method used initially to explore the legal aspects of the issue at stake is doctrinal research. At the first stage, doctrinal research has been applied as a means of extracting general principles governing the field of data security from the analysis of the relevant laws and regulations, while the process of synthesising has been based on the technique of logical deduction.

After studying the legal framework at the EU level, the method of comparative analysis has been chosen to shed light to the legal approach adopted in the jurisdictions of different Member States aiming to examine whether the way the EU Directives have been implemented to the national legislation is akin among the jurisdictions under investigation. In particular, attention has been attached to the wording of the text of each national law in order to identify any divergences in the approaches to addressing the issue of data security. After identifying the data security regulation paradigm, the need to explain the reasons, owing to which the current EU legal framework has failed to adequately deal with the problem, resulted in deconstructing the paradigm prevailing in this field. To this end, content analysis has been used to reveal both the weaknesses and the hazards of the conceptual foundations on which the legal framework has been established, which in turn led to the need for examining

the problem from different angles to be able to fully comprehend all the possible parameters. To achieve this goal, interdisciplinary research has been considered as the most suitable method since not only could it provide insights from other disciplines but also could generate a new approach to addressing the issue of data security by pointing to ways in which the gaps identified in one discipline could be covered by another discipline by integrating the various elements. Finally, the integration feature of the interdisciplinary research has been used to review the law-making process and make suggestions on how to reform the process on more realistic grounds that take into account as many factors as possible without being restricted to the legal dimensions of the problem. Therefore, applying the reform-oriented research method, suggestions are made on how to amend the philosophy underlying the law and, additionally, a new regulatory model is proposed, which introduces a new data security paradigm, aiming at overcoming the deficiencies of the current one, and hence potentially addressing the problem at stake in a more efficient manner.

2.2. Doctrinal Research

As a starting point, the method of doctrinal research has been considered as the most suitable means of conducting a critical conceptual analysis of relevant legislation in order to reveal a statement of the law relevant to the matter under investigation (Van Gestel & Micklitz, 2011). The essential features of this method involve a synthesis of legal concepts of all types (rules, principles, norms), which help explain a segment of the law as part of a larger system of law by extracting general principles form primary materials (Hutchinson & Duncan, 2013). Doctrinal method has been chosen in order to provide a systematic exposition of the principles governing the particular field of law, analyse the relationships between different approaches to addressing the problem, and explain areas of difficulty (Pearce et al., 1987). Doctrinal method is a two-part process, where the first step lies in locating the sources of the law and the second lies in interpreting and analysing the text (Hutchinson & Duncan, 2013). The research question that first emerged in the initial stage of the research process was associated with identifying which sources of the law could provide clear directions that would help unfold more thoroughly all the legal nuances of the issue at stake. First, the legal framework chosen was that of the EU, while certain elements of other jurisdictions are mentioned only when necessary to compare or contrast with the principles underlying the EU legislation. Second, although the thesis touches on issues relating to the general facets of

cyber security, and thus also studies the relevant laws, it primarily focuses on the laws regulating data protection and, more specifically, the security requirements imposed on data controllers. As a criterion of determining which laws to study was used the question of whether certain data protection laws specify any types of security obligations on data controllers. Therefore, the thesis excludes from its scope EU laws whose security requirements do not deal with the processing of personal data.

The second step, the stage of synthesising, has been based on the technique of deductive reasoning, that is, the process of reasoning based on two or more premises to reach a logically certain conclusion (Hyde, 2000). The analysis and interpretation of the relevant legal provisions have led to the conclusion that the principles underlying the EU legislation are closely intertwined with the technology neutral approach chosen by the legislator to address data protection issues emerging from the use of technologies. In order to explain the motives of lawmakers for making such a legislative choice, the concept of technological neutrality is unpacked and the rationales underlying this concept are thoroughly analysed. The ultimate goal of the synthesising stage is to answer the question of whether it is sufficient for data controllers to implement the security obligations imposed on them or there is a need to define more detailed obligations for data security. In other terms, the analysis of the technology neutral approach to addressing the data security problem aims at revealing the cases in which this approach is the most suitable as well as the cases where a technology specific approach should be preferred instead, and hence answer the question of whether the data security case falls under the first category. To this end, the analysis of the relevant national legislation of certain EU Member States has been considered as a useful tool that demonstrates whether national legislators opted for the same approach when implementing the EU laws into national laws, since any divergences in the approach adopted could provide a different perspective on the ways of handling the issue at stake. The choice of the jurisdictions to be examined has been based on the different stance of each legislator towards data protection – the flexible UK and French laws as opposed to the more detailed German law and the stricter Spanish law. In addition to the national laws of EU Member States, the data protection law of Israel was studied as an example of a recently enacted law, as opposed to the older laws of the EU Member States, because it is indicative of the additional parameters that need to be taken into account when regulating the emerging technologies; another motive for studying the Israeli law were its similar approach with the respective Spanish law. Applying the method of comparative analysis, the comparative review of the existing laws among different

jurisdictions provided interesting findings as to the debate over which approach is the most appropriate in the data security context and revealed that the technology neutral approach should not be seen as a one-way solution.

2.3. Content Research

In order to gain a more complete understanding of the conceptual bases of legal principles, and of the combined effects of the rules and procedures that touch on the area of data security, the second stage of the research process is based on the method of content research. Content research serves as a means of studying the philosophy underlying the law itself, and thus provides a critical perspective on the choices made by the legislator (Pearce et al., 1987). In particular, content analysis identifies patterns in texts and themes in bodies of documents in order to derive meaning behind the words of judicial and legislative text (Hutchinson, 2015). Hence, content research serves as a way of deconstructing the text rather than synthesising from the text as is the case of doctrinal research. Applying this method to the data security context helped unfold the concept of technological neutrality and reveal hidden aspects of this approach. Deconstructing the laws, where legislators had adopted a technology neutral approach, provided insights on the mindset underlying the EU data protection legal framework. The first step of the content analysis involved questioning the functionality of the fundamental blocks of the data protection regime by exploring the meaning lying behind the relevant legislation, which revealed a central hidden technological assumption. More specifically, the analysis of the legislation revealed that the faith in robust anonymisation, which has thoroughly infiltrated the data protection regime, is indicative of the digital technological paradigm into which the fundamental blocks of this regime are rooted. Paradigm is a shared worldview within a discipline that encompasses any underlying philosophies (Kuhn, 1996). In this regard, the technological paradigm upon which the data protection legal framework is established is sufficient as long as the paradigm remains the same. However, the fact that the law is based on a technology of a particular capability to determine whether personal data will fall within or outside the scope of the application of the data protection regime demonstrates the constraints of the legal system as it assumes a technological environment that fails to embrace emerging technologies such as social networking environments as well as the rapid expansion of Big data.

The second step of the content analysis involved studying pieces of the EU legislation that are considered to be technology neutral but are not closely linked to the data security field. The aim of this research stage was to explore how the concept of technological neutrality works in other areas of law having as an ultimate goal to identify aspects of this approach that would help answer the question of whether it is the most suitable approach to addressing the issue at stake. To this end, the EU legal framework surrounding the electronic signatures was chosen to be studied due to the technical details included in a seemingly technology neutral law. The analysis of the legislation revealed the technology favourism hidden in the text of the law since certain technologies are favoured by being afforded special assumptions, which seems to undermine the purportedly technology neutral nature of the law, and is indicative of the legislator's failure to achieve the goal of drafting laws that do not favour any particular technological design. Concerns stemming from the technology neutral laws in the field of surveillance and data processing activities called for the need to examine the relevant legislation in order to explore the consequences of making such a legislative choice in this area of law. Content analysis of the surveillance legal framework identified the potential hazards hidden in adopting a technology neutral approach, which means that technical considerations with severe implications are left to secondary legislation and excessive powers are granted without due process, thus jeopardising individuals' fundamental rights.

2.4. Reform-oriented Research

The deconstruction of the concept of technological neutrality demonstrated that the widely accepted view shared among legislators, according to which the enactment of technology neutral laws is the most suitable means of guaranteeing the sustainability of laws while granting flexibility to the private sector, often leads them either to ignore or fail to foresee the implications and the hazards involved in certain technology neutral laws. In other words, dogmatically embracing the principle of technological neutrality often carries the risk of abandoning legal certainty. Nevertheless, despite the increasing number of re-identification incidents, there has not so far been any amendment in the law addressing the weaknesses and gaps identified in the current legal framework on the security obligations of data controllers. At this stage of the research process the method of reform-oriented research has been used to leverage the outcome of the content analysis presented above. Contrary to doctrinal research, which identifies and analyses the current laws, reform-oriented research goes beyond

description, analysis and critique, and evaluates the adequacy of existing laws while also suggesting ways the law could be amended or how the philosophy underlying the law as well as the administration of the law could be improved in order to bring the law into line with current conditions and ensure that it meets current needs (Pearce et al., 1987; Weisbrot, 2005).

Applying the reform-oriented research method, the thesis makes recommendations in terms of the legislative model that would better fit in the context of data security and proposes a new regulatory model, which challenges the conventional law-making process and calls for a shift in the data security paradigm. In spite of the fact that a paradigm is considered to be a shared frame of conference among researchers, new revelations can make it change and this is occurring within the discipline of law (Kuhn, 1996). Once a paradigm changes, new regulatory models are required and, consequently, lawmakers need to find new ways to regain the lost balance. For example, the once-prevalent view of law as being objective and neutral, or else law as being 'what is', should be replaced by the view of law as being 'what could be' or 'what should be' (Hutchinson, 2015). Hence, based on the findings of the content analysis, the thesis suggests that the static approach to the definition of personal data adopted by the current law be replaced by a dynamic approach, where the line between personal and non-personal data is not fixed but depends on the constantly changing technologies, and the law recognises that identifiability cannot be determined a priori but is rather driven by context. Another parameter that has been taken into account when developing the proposed model is the principle that laws should refer to a certain technology in such detail that allows for the benefits of technological specificity but also generally enough to prevent the need to revisit the law frequently.

2.5. Interdisciplinary Research

Reform-oriented research has the potential to be interdisciplinary in its methods and, as a matter of fact, legal scholars are increasingly infusing evidence from other disciplines (such as statistics, comparative perspectives, social science) into their reasoning to bolster their reform recommendations (Weisbrot, 2005; Hutchinson, 2015). Interdisciplinary research is a process of answering a question or solving a problem that is too broad or complex to be dealt adequately by a single discipline or profession (Klein & Newell, 1997). As the foregoing

analysis has demonstrated, data security is a complicated matter with various dimensions that need to be taken into account when making recommendations on how to address the problem more effectively. Therefore, merely studying the legal facets of an issue that is closely intertwined with the use of technologies without considering the technical aspects of the issue would result in a fruitless attempt to suggest a workable solution. Furthermore, as it has been explained above, one of the reasons why the current legal framework fails to address the problem is associated with the lack of flexible policies that are tailored to the volatile nature of new technologies, which calls for the need to introduce dynamic security concepts able to allow improvements that can best meet emerging threats. In order to develop a regulatory model that would overcome the challenges posed by the current legal framework and incentivise the private sector to implement robust security measures, it would be essential to integrate industry's practices into the law-making process as a potential means of striking the right balance between regulation and innovation. For the aforementioned reasons, the disciplines of Computer Science and Managerial Economics have been chosen to enrich the legal analysis. More specifically, the discipline of Computer Science could deal with the technical considerations, such as the factors that determine the re-identification risk level, the security measures that need to be applied depending on the security level required, and the security objectives that need to achieved in each case. As far as the discipline of Managerial Economics is concerned, it could provide insights on the changes that need to be made in the law-making process in order to accommodate the rapid pace of technological advances. In particular, studying decision-making models used in Economics could help materialise the 'post-bureaucratic' vision of the law, in which regulators govern through the use of auditing and continuous adaptation to diverse and changing environments, and thus reframe the elements currently forming the regulatory decision-making process.

After drawing on disciplinary perspectives, the last stage of the research process involves the element of integration; at this stage, the different perspectives are integrated through construction of a more comprehensive perspective producing an interdisciplinary understanding of the problem (Repko, 2011). The potential of the proposed regulatory model depends on the way legal factors interact with technical and economic factors. In other terms, a new regulatory model would require understanding different business models and emerging technologies, monitoring of market mechanisms and practices, and utilisation of new tools such as management and algorithmic solutions. The aim of this stage is to first identify the conflicts between disciplinary insights and then discover or create a common ground. For

example, the ‘all-or-nothing’ legal approach to the definition of personal data appears to contradict with the concept of negligibility prevailing in engineering, which suggests that in practice there is no zero-risk situation, and thereby data should be considered anonymised in case of a low re-identification risk. Hence, the construction of the two different perspectives should establish a common understanding of the notion of the re-identification risk between technologists and lawmakers. In order to achieve an optimal balance between protecting personal data and promoting private sector’s interests, the proposed regulatory model recognises the existence of different degrees of data identifiability, and proposes a holistic, dynamic, hybrid and layered approach to data security that systematically tailors the security obligations of data controllers to the level of re-identification risk involved in data processing activities.

CHAPTER 3: Should the EU Legal Framework Surrounding Data Controllers' Security Obligations Be Technology Neutral?

3.1. Introduction

The emergence of information and communications technologies (ICTs) has prompted the demand for the new regulatory regimes to adopt a level of abstraction and thus abstain from targeting specific technologies to be regulated or used for the implementation of certain provisions. Technology neutrality has long been held up as a guiding principle for the regulation of technology in the field of ICT regulation and continues to be a pervasive concept in this field. Since 2011, technology neutrality has also been recognised as a key principle for Internet policy (OECD, 2011). Although the principle of technology neutrality is widely accepted as it often figures in policy documents and regulatory instruments, its meaning is not always clearly understood. This chapter starts with unpacking the concept of technology neutrality and examines the reasons that lead lawmakers to apply this principle on technology-related regulation. In order to clarify the substance of this principle, focus will be placed on the rationales underlying technology neutrality and on its utility in different contexts. The analysis of the repercussions that a blind adoption of the technology neutral approach might have in the law-making process is indicative of the need for regulators to weigh up the pros and cons of this principle and opt for technology specific regulation if deemed as more suitable, especially in the cases when the former could potentially impact upon fundamental human rights. In an attempt to identify which approach should be preferred in each case, certain criteria relating to the purpose, context and means of regulation are presented, while different ways to strike a right balance between the conflicting needs for sustainability and for legal certainty are suggested.

The purpose of this chapter is to examine whether the EU data protection legal framework is formulated in a technology neutral way and, in particular, whether data protection laws specify any types of security obligations imposed on data controllers. In other words, this chapter aims to answer the question of whether it is enough for data controllers to implement the security obligations incumbent on them or there is a need to define more detailed obligations for data security. To this end, the second part briefly explains the reasons for

which technology neutrality has been established as one of the general principles underpinning the EU regulatory framework, before progressing with the analysis of the EU legislation. The EU Directives stipulating security obligations are presented in order to examine whether the wording of the relevant provisions contains technology neutral elements. For the same purpose, the UK, German and French Acts transposing the DPD into national legislation are studied, while special emphasis is placed on the Spanish Act due to its different regulatory approach. The recently enacted data protection law of Israel is also presented owing to its similarities to the Spanish legislation. In order to clarify whether the proposed reform of the data protection framework will bring significant changes to the regime of data controllers' obligations, the GDPR is analysed and the extent to which the concept of 'data protection by design' included in the Regulation violates the principle of technology neutrality is examined.

3.2. Technology Neutrality: Overview

3.2.1. Definition of technology neutrality

Technology neutrality can be seen as a requirement that ICT legislation should not focus on specific forms of technologies; in other words, the same regulatory principles must apply irrespective of the technologies used. What is stressed by the use of this principle is that the goal of regulation is to regulate the effects or the functions of the technologies targeted but not technology itself so as to achieve equivalence between offline and online regulation; the fundamental rules should be the same for an online technology activity as for its equivalent offline (Koops, 2006). This approach has been adopted by the UK e-Principles, according to which "regulation should be technology neutral *in its effects*. The effects of the offline and on-line regulatory environments […] should be as similar as possible".[6] In this context, technology neutral laws should provide for the necessary *performance standards* as opposed to *design standards*. Performance standards – standards that describe the result to be achieved without imposing a given technology – are considered to be more efficient as they leave the regulated entities free to choose which technology is the most appropriate in order to achieve the outcome specified in the standards. On the contrary, in the case of design standards, the

[6] E-Policy Principles - A policymakers guide to the Internet, available at:http://webarchive.nationalarchives.gov.uk/20040722012351/e-government.cabinetoffice.gov.uk/assetRoot/04/00/60/79/04006079.pdf.

choice of technology to be adopted is made by the regulator and thus this type of legislation is called technology specific since it incorporates technological choices (Hemenway, 1980; Breyer, 1980). The hazards of technology specific legislation have as a result policy makers to resort to the principle of technology neutrality.

Technology neutrality principle performs a twofold function in the current regulatory framework; the rationales underlying this principle are related to the need for sustainability and innovation (Ali, 2009; Hildebrandt & Tielemans, 2013). Under the rationale of sustainability, which refers to the technique of lawmaking, regulation should be flexible and time-proof so as not to require over-frequent revision to cope with technological changes. Given the rapid pace of ICT development, specifying the technologies that should be implemented would result in regulation being outdated and thus inefficient in a short period of time (European Commission, 1999). The rationale of innovation contains a negative requirement aiming to avoid certain side effects of technology specific regulation. Incorporating design standards in the regulation means that the regulator would favour certain technological designs against others thus pushing the market toward a particular structure, which would harm competition and would prevent the existence of a highly dynamic market (Gervais, 2005). It is also likely that the choice of the regulator with respect to the design standards would be subject to the influence of strong industry players having the resources to lobby for a particular technological choice (Samuelson, 2000).

3.2.2. Sustainability: making laws easily adaptable to technological changes

Given that law always lags behind technology, the sustainability of laws is considered to be of vital significance in ICT regulation, where the technological advances are more likely to render the law obsolete compared to non-technological types of regulation. The sustainability of ICT laws was highlighted many years ago, in 1997, in the EU Bonn Ministerial Conference, which stated: "in view of the speed at which new technologies are developing, they will strive to frame regulations which are technology-neutral, whilst bearing in mind the need to avoid unnecessary regulation".[7] In order to prevent laws that target technology from becoming out of date sooner than expected, the sustainability requirement dictates the need to enact legislation at the right level of abstraction so as not to require continuous adaptation to

[7] Declaration of the European Union Ministers, Global Information Networks: Realising the potential, July 6-8, 1997, Bonn, available at http://cordis.europa.eu/news/rcn/8627_en.html.

emerging technologies. The rapid advance of technology calls for future-proof legislation able to cover future technological developments or, in other words, able to accept any technology that the market could develop in the near future (Reed, 2007). The risks generated by technology specific regulation are associated with the lawmaking procedure as well as the need for legal certainty. First, the process of constructing legislative acts is time-consuming and cumbersome and thus enacting new legislation to address issues emerging from new technologies would be not only difficult but also ineffective; by the time a new law would be enacted, the technology targeted would be regarded obsolete. Second, given that the role of law lies in mandating individuals to behave in a certain way, changing the legal norms at such a speed that could no longer provide for legitimate expectations would contrast with the need for legal certainty (Hildebrandt & Tielemans, 2013).

On the other hand, however, legal certainty is challenged by the unforeseen impacts of the future usage of the technologies targeted or of new technologies with similar effects, especially in cases that the legislator does not fully comprehend the technology it targets. The sustainability paradox consists in the fact that law aims at regulating technology but it should not be amended when future technologies emerge, whereas emerging technologies can even alter the scope of application of existing legal norms (Taleb, 2007). Technology neutrality requires laws to be formulated in such a way that they can predict the unpredictable but at the same time they must not be over/under-inclusive. It might be necessary to reconsider the meaning of the sustainability requirement itself in the digital era, when the rapid technological advance results in technology-related laws being unsustainable in a shorter period of time. Another risk of enacting sustainable legislation is that the interpretation of the law may diverge for different technologies over the years thus leading to adverse results such as unintended technology specificity (Koops, 2006). There are also cases where technology neutral laws are formulated in such a way that end up being meaningless. The following comment of the Earl of Northesk on the UK 2000 Regulation of Investigatory Powers Act is illustrative of this point: "One of the many difficulties I have with the Bill is that, in its strident efforts to be technology neutral, it often conveys the impression that either it is ignorant of the way in which current technology operates, or pretends that there is no technology at all."[8]

[8] Hansard, House of Lords 28th June, 2000 (Committee Stage), Column 1012, available at http://www.publications.parliament.uk/pa/ld199900/ldhansrd/vo000628/text/00628-28.htm.

3.2.3. Innovation: making laws open to market entries

The fact that the legislator is unable to foresee the technological designs that will be developed in the future means that any technology specific legislation would only address technologies that are already on the market or those that the legislator could predict at the moment of the enactment of the law. This would result in an unfair competitive disadvantage for new technologies as barriers would be created to market entries. However, it is crucial that law abstains from imposing unnecessary constraints on the development of new technologies and business models. One of the objectives of technology neutrality is to prevent law from unduly discriminating against certain technologies and influencing the users in order to benefit certain developers to the detriment of their competitors. The reason is that unjustified discrimination would result in interference with the market dynamics of competing technologies and would stifle innovation (Shelanski, 2013; Lovells, 2014).

In an attempt to promote competition and remove barriers from the internal market, in the 1999 Communications Review, the European Commission referred to technology neutrality declaring that "rules should neither impose, nor discriminate in favour of the use of a particular type of technology" (European Commission, 1999). Similar is the objective of the Article 14 of the e-Privacy Directive, according to which no mandatory requirements regarding specific technical features should be imposed on terminal or other electronic communication equipment "which could impede the placing of equipment on the market and the free circulation of such equipment in and between Member States".[9] There are, however, two exceptions to the general principle of technology neutrality in terms of promoting innovation. First, in the case that certain technologies infringe by default upon fundamental rights such as the right to privacy and data protection, it is acceptable that law discriminates against them (Kannecke & Körber, 2008). Second, technology specific law is preferable when there is a significant difference between the technologies targeted in terms of their functions or effects; within the same type of technology, however, they should be subject to the same legal conditions. At this point, it is worth noting that the innovation objective of technology neutrality seems to reinforce the adoption of self-regulation in the ICT area as a more effective means of fostering technological development (Lovells, 2014).

[9] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector ("e-Privacy Directive").

3.2.4. Technology specific law

As indicated above, the reasons that make the principle of technology neutrality be deemed as a better option in the ICT arena can easily be questioned and therefore this legislative technique needs to be qualified in every single lawmaking case. To begin with, the aim of equivalence between online and offline technology activities sometimes appears to be an unrealistic aim. The unique features of the online environment render these activities different from their offline counterparts to such an extent that it is hard even to define what online and offline equivalence might possibly mean; instead, a different approach often needs to be adopted with respect to online activities to create similar effects of legal protection offline. Moreover, technology neutrality can also produce undesirable consequences, such as ex-ante regulation and legal uncertainty. One of the assumptions underlying the technology neutral approach is that regulation will be sustainable and thus able to deal with technological changes; however, this has not always proved to be true in practice. Given that technological development reflects the socioeconomic changes, it is quite difficult for the legislature to foresee those changes and enact laws that would also address them. Often, the failure of regulators to achieve the original legislative aims derives from changes in the underlying business models. A hazard involved in technology neutral regulation is that legislators might be tempted into regulating prospectively before thoroughly comprehending the potential uses of a new technology and what issues the regulation will actually need to solve (Reed, 2007). Yet, even if all the aspects of a technology are understood, it is questionable whether any technology neutral law can be proof against unanticipated changes to that technology.

A further effect of technology neutral regulation is legal uncertainty. Although the requirement for legal certainty means that legislation should not unduly discriminate, in which case technology neutrality is the most optimal choice, there also lurks the danger of regulation whose meaning is completely vague. An unsuccessful attempt to achieve technology neutrality is likely to result in insufficiently clear regulation whose application to technology is often “a matter of guesswork” (Escudero-Pascual & Hosein, 2004). The need for legal certainty could be a reason for technology specific regulation called for instead. As stated in the Dutch policy memorandum Legislation for the Electronic Highways of 1998: “A starting point is that rules are technology-neutral. In formulating technology-independent

rules, however, it should be considered whether these guarantee sufficient legal security".[10] In contrast to neutrality, which is sometimes subject to various interpretations as to the scope of the regulation, specificity creates substantial certainty with respect to which fields of activities are covered by that regulation and what needs to be done to comply with it. Often, technology specific regulation is rejected at the outset because technology neutrality is so widely accepted that it is usually seen as the only suitable way to cope with legal issues in the ICT area (Reed, 2007). There are occasions, however, when technology specific drafting can be proved as a more effective technique. Such is the case of technology designs that threaten fundamental human rights and therefore call for specific provisions to retain the substance of the legal right they support (Kannecke & Körber, 2008). Specific laws should also be put in place to regulate the use of technologies that fundamentally alter the nature of the activity to be controlled – when the effects of a technology to be mandated or prohibited are made different in kind by the means adopted by the regulator (Reed, 2007). Finally, although technology neutrality is the default choice, in order to achieve neutrality, sometimes technology specific legislation may be required as is the case of laws treating different types of technologies.

Technology specificity is not, of course, without limitations. In their attempt to produce accurate and detailed laws, regulators may end up adopting an overly specific approach, which cannot accomplish its target precisely because of a high degree of specificity (Hildebrandt & Tielemans, 2013). Furthermore, trying to ensure that regulation keeps pace with technological changes, technology specific legislation challenges the traditional lawmaking system as it forces the legislator to reconsider the regulation regularly. Given the cost in legislative time and effort required to keep technology specific regulation up to date, especially when recurrent updates are needed, it is doubtful whether such regulation would cope adequately with the rapid pace of technological development. As mentioned above, however, the same problem seems to be a limitation of technology neutral legislation too.

3.2.5. When is technology neutral regulation suitable?

Koops (Koops, 2006) has listed certain criteria relevant to the applicability of the technology neutrality principle dividing them into three categories. The first set of criteria related to the

[10] LEH Memorandum 1998, p.12.

purpose of the regulation involve questions, such as what the goal of regulation at issue is and whether it is desirable to control the use and development of technology or it would be better to leave it alone without any legal intervention. The next two criteria pertaining to the same category refer to the notion of legal certainty. Despite the fact that legal certainty is undoubtedly a general requirement of regulation, the level of legal certainty called for may vary greatly depending on factors such as the scope of regulation and its subject matter. The need for legal certainty also determines how urgent the need for regulation is. In case of emerging new technologies, technology specific legislation is more likely to provide the level of legal certainty needed, especially when such a technology impacts on the legal protection of basic rights. If, on the other hand, there is no urgent need for providing legal certainty, high-level legislation is the most suitable option as it is formulated in such a way that meets the sustainability requirement.

Moving on to the second category – criteria related to the context of regulation - Koops refers to 'technological turbulence' to point out that with technologies developing at a high speed, such as mobile communications and computer processors, lower-level forms of regulation should be chosen due to their ability to adapt more easily to rapid technological advances than formal regulation. The scope of interpretation, that is, the extent to which legislation leaves room for interpretation, also influences the degree of technology neutrality required. The issue of supervision and enforcement is the last criterion of this category; in the cases when regulation is difficult to be enforced in practice, technology specific legislation focusing on means rather than effects should be put in place. As far as the third category is concerned, the criteria related to the means of regulation are closely intertwined; the degree of sustainability called for determines the level of regulation required. Technology neutrality appears to be suitable for high-level regulation, such as constitutional regulation, as it is formulated in a more abstract way, whereas low-level regulation can be more detailed and thus technology specific.

Although drafting technology neutral laws appears to be the most widely adopted approach in the lawmaking process, it is crucial for regulators to know when they should avoid technology neutrality. Instructive in this respect is the list of criteria articulated in the first Dutch memorandum on ICT regulation,[11] according to which technology neutral laws are not

[11] LEH Memorandum 1998, p.14.

suitable in the following cases: as a definition of a scope of regulation; when there is a need for understanding complex technologies affecting legal rights; when the nature of legal subjects' rights and obligations cannot be adequately described by technology neutral rules; when there is the risk of government violating those rights. Quoting Ian Hosein & Alberto Escudero Pascual (Hosein & Pascual, 2004) sheds more light to the hazards of drafting technology neutral regulation without taking into account all the parameters. "Attempts to be technology-neutral should be interrogated, lest in our blindness we reduce democratic protections and oversight under the deterministic veil of progress."

3.2.6. Technology neutrality vs. technology specificity

When regulating new developments in ICT, lawmakers are responsible for making the most suitable choice between technology neutral and technology specific approaches. Even though neutrality is deemed as the most appropriate technique for regulating technological areas, lawmakers ought to consider whether technology specific approach would produce better regulation instead. As the above analysis has demonstrated, the crucial criterion for adopting the most effective approach is the primary requirement of regulation to provide sufficient legal certainty. However, the principle of technology neutrality often conflicts with the need for clearly specified goals upon which regulation has to be based; the requirement for legal certainty may call for a certain level of technology specificity. More specifically, the sustainability objective of the neutrality principle appears to be the most problematic. On the one hand, regulation that abstracts too much away from technology results in vague laws that do not sufficiently provide legal certainty. On the other hand, legislation which contains too much detail in terms of explicating technologies risk being inflexible and restricted to certain technologies without encompassing different but relevant technologies.

As it is apparent, opting for a particular legislative technique entails dealing with a trade-off between sustainability and legal certainty; the one calling for flexibility and the other for predictability (Koops, 2006). The most effective way to strike the right balance between these requirements is combining technology neutral and technology specific approaches. Laws should be formulated in a sustainable way providing at the same time clear guidance as to the aims and the rationale of the regulation by explaining what technologies they cover and why. In particular, technology neutral laws can establish a general regime covering all different technologies by defining the performance standards, that is, the desired outcome, and a

regime for specific technologies by providing examples of technologies that will satisfy the output described in the standards (Hemenway, 1980). Another way to achieve this balance is enacting multi-level legislation by laying down the basic requirements in high-level legislation and filling it in with lower-level forms of regulation, which will elaborate this requirement in more technical detail (Koops, 2006). Self-regulation can also be included based on guidelines with no legal effect, which would allow stakeholders to coordinate their behaviour in the market.

3.3. Technology Neutrality in the EU Regulatory Framework

3.3.1. Technology neutrality in the EU regulatory framework: Overview

According to the general principles underpinning the EU regulatory framework, regulation should be based on clearly defined objectives; be the minimum necessary to meet those objectives; be enforced as closely as practicable to the activities being regulated; further enhance legal certainty in a dynamic market; and aim to be technologically neutral (European Commission, 1999). In 2002, technology neutrality was also recognised as one of the key regulation principles in the field of electronic communications; the principle was first introduced by the Framework Directive[12] and was reinforced with the revised EU telecoms legislation in 2009.[13] The principle of technology neutrality is defined in Recital 18 of the Framework Directive, which refers to the requirement for Member States to ensure that "national regulatory authorities take the utmost account of the desirability of making regulation technologically neutral, that is to say that it neither imposes nor discriminates in favour of the use of a particular type of technology". Technology neutrality is also mentioned in Article 8(1) of the Directive as an obligation imposed on national regulatory authorities when carrying out regulatory tasks "designed to ensure effective competition". As mentioned above, Article 14 of the e-Privacy Directive prohibits the imposition of mandatory requirements for specific technical features that would hamper the placing on the market and

[12] Directive 2002/21/EC of the European Parliament and of the Council of 7 March 2002 on a common regulatory framework for electronic communications networks and services ("Framework Directive").

[13] Directive 2009/140/EC of the European Parliament and of the Council of 25 November 2009 amending Directives 2002/21/EC on a common regulatory framework for electronic communications networks and services, 2002/19/EC on access to, and interconnection of, electronic communications networks and associated facilities, and 2002/20/EC on the authorisation of electronic communications networks and services ("Telecoms Directive").

the free circulation of any electronic communication equipment. Recital 46 of the e-Privacy Directive explains that the background of this Article is the technology neutrality of the DPD which "covers any form of processing of personal data regardless of the technology used", while the Recital highlights that the existence of specific rules for electronic communications services might not facilitate the protection of personal data and privacy "in a technologically neutral way". Finally, Recital 49 of the Tecoms Directive stresses the need for technology neutral laws, that is to say laws that can adapt rapidly to the high level of technological innovation and the highly dynamic markets in the electronic communications sector.

In order to encourage competition in electronic communications markets, the EU regulatory framework adopts a technology neutral approach, which allows for more flexibility for regulators to make account of market needs and thus better respond to the dynamic and largely unpredictable communications markets (Koenig, 2009; Kerikmäe, 2014). The introduction of the concept of technology neutrality should be examined in the light of increasing convergence of technologies, networks and services and the transition from monopoly to competition, which called for a unified approach to regulation. In contrast to the previous regulatory framework, where each type of network was subject to separate sets of rules, the current framework requires regulators to act in a technologically neutral way applying the same principles of market analysis and remedies to all kinds of electronic communication networks and services (Savin, 2014). It has been argued that another reason behind the adoption of a technology neutral approach is the fact that this approach empowers regulators to push the market toward self-regulatory or co-regulatory solutions, which can provide guidance to regulated entities and perhaps be more effective than command-and-control regulations (Halftech, 2008).

3.3.2. EU Directives imposing security obligations on data controllers

3.3.2.1. Need to implement "appropriate technical and organisational measures"

The technology neutral approach adopted in the EU legal framework is reflected in the wording of the Directives that impose security requirements on data controllers. By virtue of Article 17 of the DPD, Article 13a of the Framework Directive and Article 4 of the e-Privacy

Directive (as amended by Directive 2009/136)[14], data controllers are required to take "appropriate technical and organisational measures" in order to safeguard the personal data they process. In order to be considered as appropriate, security measures must meet the following criteria: they should involve state-of-the-art technologies; the investments should be proportional to the potential of the controller; they should take into account the nature of the data (e.g. stricter for financial data, health data etc. compared to mere contact data); they should be in line with the potential risks represented by the processing (e.g. financial institutes are a common target for hackers). There is a large overlap between security measures that have to be taken as EU law establishes a general obligation for adequate security focusing on the need to implement the aforementioned measures, which is mentioned in all the security articles. Based on the premise that there is no one-size-fits-all solution in data security, EU law encourages organisations processing personal data to adopt a risk-based approach and thus determine the necessary level of security depending on their own circumstances and risks. It is clear that all three Directives are formulated in a technology neutral way as there are no specific requirements in terms of the technical features of the security techniques used. In order to inform all the relevant stakeholders of the best available means of implementation of the security articles, ENISA and Article 29 WP regularly issue guidance specifying the technical implementing measures to be adopted.[15]

3.3.2.2. The e-Privacy Directive 2002/58/EC (as amended by Directive 2009/136/EC)

The principles set out in the DPD are translated into specific rules for the telecommunications sector in the e-Privacy Directive. As opposed to the Framework Directive, which applies to both providers of publically available electronic communications networks and services, the security requirements stipulated in Article 4 are limited to service providers. As this term is not defined in the e-Privacy Directive, the definition provided in Framework Directive

[14] Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws.

[15] See e.g. ENISA, Technical Guideline on Security Measures: Technical guidance on the security measures in Article 13a, Version 2.0, October 2014, available at http://www.enisa.europa.eu/activities/Resilience-and-CIIP/Incidents-reporting/technical-guideline-on-minimum-security-measures; Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques, adopted on 10 April 2014, available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

Article 2(c) is applicable in this case.[16] According to this definition, "electronic communications service" is "a service normally provided for remuneration which consists wholly or mainly in the conveyance of signals on electronic communications networks." In this sense, the term includes "telecommunications services and transmission services in networks used for broadcasting" but services "providing, or exercising editorial control over, content", therefore content providers, are excluded. A distinction worth mentioning is that between the provision of electronic communications services and providers of online services. The fact that the former has to consist in the conveyance of signals differentiates it from services which merely initiate the conveyance. From a technical perspective, the conveyance of the signals takes place on the data-link and network layers of the TCP/IP protocol suite and therefore as electronic communications services should be understood services provided on the first two layers of the abovementioned networking model. Hence, the scope of application of e-Privacy Directive Article 4 covers telephone and Internet access providers as well as Internet backbone providers. On the contrary, providers such as hosting providers or email service providers do not perform the actual conveyance of the signals and thus do not fall under the definition of providers of electronic communications service (Feiler, 2011).

The security requirements imposed on data controllers in the DPD for the purpose of improving data security are rephrased in the context of the e-Privacy Directive, where service providers play the role of data controllers as defined in the DPD. As far as companies (e.g. ISPs) provide publically available communications services (e.g. via public websites) over public networks (e.g. Internet), there are similar legal security obligations to those described in the DPD. Article 4 requires providers of publically available communications services to "take appropriate and organisational measures to safeguard the security of their services". In the case that a service provider relies on a provider of a public communications network to provide its service, the former must take security measures "in conjunction" with the latter. In terms of what measures can be considered "appropriate", Article 4 provides similar criteria with the DPD stating that "these measures shall ensure a level of security appropriate to the risk presented" taking into account "the state of the art and the cost of their implementation". More detailed security obligations with regard to the appropriate security measures that need to be taken were placed by the Directive 2009/136/EC. In particular, pursuant to Article

[16] See e-Privacy Directive Article 2.

4(1)(a) service providers must ensure the authorised access to personal data as well as the implementation of a security policy with respect to the processing of personal data. Service providers are obliged to protect personal data "against accidental or unlawful destruction, accidental loss or alteration, and unauthorised or unlawful storage, processing, access or disclosure". What is particularly significant is the fact that Article 4(1)(a) explicitly states that the security obligations of service providers apply to personal data both stored and in transit. The same article imposes on national regulatory authorities the responsibility to audit the measures taken and issue recommendations regarding the level of security that needs to be achieved by these measures.

3.3.2.3. Legal requirement for the adoption of encryption mechanisms

Despite the fact that the EU data protection legislation is written in a technologically neutral way as it avoids being prescriptive about the technologies that need to be adopted for security purposes, it could be argued that, whether expressly or by implication, data protection laws in the EU context give rise to a requirement for the adoption of encryption technologies. First, it is clear that Article 4(3) of the e-Privacy Directive provides an explicit exemption from the breach notification requirements in case the data are "unintelligible to any person who is not authorised to access it". In other words, service providers are not obliged to notify the affected individuals of a security breach as long as they are able to prove to the competent authority that they had applied technological measures that can prevent any unauthorised third party from "reading" the data concerned by the security breach. The most common means of rendering data unintelligible, as required by the Directive, is the use of technologies that encode information in such a way that only authorised parties can read it, that is, encryption technologies. It should be noted, however, that the implementation of encryption exempts service providers from the breach notification requirement on the condition that data has been effectively encrypted, which means that the encryption key has not been compromised and it cannot otherwise be ascertained by available technological means. In more general terms, it can be argued that the adoption of encryption technologies as a legal obligation was established prior to the e-Privacy Directive. As mentioned above, Article 17 of the DPD requires data controllers to take appropriate security measures having regard, amongst others, to "the state of the art"; encryption is a security technique that falls within the state of technological development and consequently should be deployed in appropriate

situations.[17] In this regard, the DPD has created a legal environment for encryption (Room, 2014).

Likewise, the national data protection regulatory authorities of the EU Member-States seem to expect organisations to implement encryption techniques to protect the data they process especially in cases when personal data are highly sensitive or confidential. The example of the regulatory authority in charge of enforcing data protection laws in the UK – the Information Commissioner's Office (ICO) – is illustrative of such an expectation. The ICO's regulatory guidance on best security practices seems to mandate the use of encryption technologies; regulatory enforcement action may be pursued against data controllers in case of a data breach "where encryption software has not been used to protect the data".[18] In the ICO's Practical Guide to IT Security, encryption is highlighted as "a means of ensuring that data can only be accessed by authorised users", while the same technique is also described as an important "first step" that businesses must consider when assessing their data processing operations.[19] Moving towards the same direction, the French regulatory authority – Commission Nationale de l'Informatique et des Libertés (CNIL) – has stressed the significance of using encryption technologies, such as encrypted links like "https" for electronic exchanges of data as well for the protection of data at rest in the Cloud. The CNIL clearly defines encryption as an "appropriate protection measure", which can exempt a service provider from the breach notification requirement, and highly encourages companies to use encryption as the risks to an individual are limited when "there is no possibility for an (individual file) being opened without prior decryption with a confidential password that has not been hacked".[20]

[17]*See* e.g. ENISA, 2013, Recommended cryptographic measures: Securing personal data, September 20, available at: https://www.enisa.europa.eu/activities/identity-and-trust/library/deliverables/recommended-cryptographic-measures-securing-personal-data.

[18]*See* e.g. ICO, Encryption, available at https://ico.org.uk/for-organisations/encryption/; ICO's Practical Guide to IT Security, April 2012, available at https://ico.org.uk/media/for-organisations/documents/1575/it_security_practical_guide.pdf.

[19]*See* ICO's Guide to Data Protection - Information Security (Principle 7), available at https://ico.org.uk/for-organisations/guide-to-data-protection/principle-7-security/.

[20]*See* The CNIL's Guide: Security of Personal Data, 2010, available at http://www.cnil.fr/fileadmin/documents/en/Guide_Security_of_Personal_Data-2010.pdf; CNIL (2012) Measures for the Privacy Risk Treatment, available at http://www.cnil.fr/fileadmin/documents/en/CNIL-ManagingPrivacyRisks-Measures.pdf.

3.3.2.4. Divergences in the EU Member-States' national laws

3.3.2.4.1. Debate over the technology neutral approach

The different approaches to drafting data security legislation are reflected on the divergences of the national laws implementing the DPD; for instance, German, France and Spain have stricter data security laws in place whereas the UK has more lenient laws.[21] This differentiation has also been apparent in the reform process of the data protection legislation and raised a debate over the technology neutral character of the security requirements laid down in the GDPR. On the one hand, the UK favoured a technology neutral approach which would allow great scope for flexibility pointing out the potential dangers stemming from a detailed specification of security criteria and mechanisms, and highlighting that technological neutrality in the security sphere has been one of the widely acknowledged strengths of the DPD (ICO, 2013). On the other hand, Spain expressed concerns that a technology neutral approach would not provide adequate solutions for certain challenges emerging from future technologies, such as those presented by cloud computing or the transfer of personal data over the Internet (Agencia Española de Protección de Datos, 2013), while Germany stressed that risk-based approaches should only refer to how obligations will be met and compliance requirements ought not to be relativized (Der Hessische Datenschutzbeauftragte, 2015). In particular, the UK ICO raised objections regarding the wording "state of the art" as a criterion for data controllers' and processors' liability on the grounds that their liability should be determined based on the accepted industry standards for breaches resulting from technical failings rather than on whether security measures were "state of the art" (ICO, 2013). As far as the German position is concerned, based on the opinion that data protection legislation should define goals on which the data protection measures to be oriented, the German Data Protection Commissioners stated that the classic goals of data security as set out in the Council and Commission text, that is, the goals of availability, integrity and confidentiality, could not adequately protect personal data and therefore the goals of "non-linkability, transparency and the ability to intervene" should be added (Der Hessische

[21]Bundesdatenschutzgesetz (BDSG) (promulgated on 14 January 2003 and amended by Article 1 of the Act of 14 August 2009) Section 9; Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés (Loi Informatique et Libertés) as modified to implement Directive 95/46/EC on data protection, Article 34; Ley Orgánica 15/1999 de Protección de Datos de Carácter Personal (LOPD) as supplemented by the Data Protection Regulation (Real Decreto 1720/2007 (RLOPD), Articles 89-114; UK Data Protection Act 1998, Schedule 1, Part II, para 9.

Datenschutzbeauftragte, 2015). The text proposed by the European Parliament goes further and follows a layered approach, according to which additional security measures should be applied to the processing of sensitive personal data in order to "ensure situational awareness of risks" (EDPS, 2015). Moreover, the European Parliament provides detailed guidance with respect to the security measures that should be deployed adopting the same wording as that of Art. 4 (1)(a) of the e-Privacy Directive: "the measures (…) shall at least ensure that personal data can be accessed only by authorised personnel for legally authorised purposes, protect personal data stored or transmitted against accidental or unlawful destruction, accidental loss or alteration, and unauthorised or unlawful storage, processing, access or disclosure, and ensure the implementation of a security policy with respect to the processing of personal data".

3.3.2.4.2. UK Data Protection Act

The principle data protection legislation is the Data Protection Act 1998 (DPA), which took effect in 2000 and implements into UK law the requirements of the EU DPD. The DPA requires data controllers to take "appropriate technical and organisational measures" to keep personal data safe and secure. Specific standards are not stipulated by law or binding guidance. The Seventh Data Protection Principle provides at Part I of Schedule 1 to the Act that:

"Appropriate technical and organisational measures shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data".

Paragraph 9 of Part II of Schedule 1 to the Act further provides that:

"Having regard to the state of technological development and the cost of implementing any measures, the measures must ensure a level of security appropriate to –
(a) the harm that might result from such unauthorised or unlawful processing or accidental loss, destruction or damage as are mentioned in the seventh principle, and
(b) the nature of the data to be protected."

3.3.2.4.3. German Federal Data Protection Act

Section 9: Technical and organisational measures

Public and private bodies which collect, process or use personal data on their own behalf or on behalf of others shall take the necessary technical and organizational measures to ensure the implementation of the provisions of this Act, especially the requirements listed in the Annex to this Act. Measures shall be necessary only if the effort required is in reasonable proportion to the desired purpose of protection.

Annex to Section 9

Where personal data are processed or used in automated form, the internal organisation of authorities or enterprises is to be such that it meets the specific requirements of data protection. In particular, measures suited to the type of personal data or categories of data to be protected shall be taken:

1. to prevent unauthorised persons from gaining access to data processing systems for processing or using personal data (access control),

2. to prevent data processing systems from being used without authorisation (access control),

3. to ensure that persons authorized to use a data processing system have access only to those data they are authorised to access, and that personal data cannot be read, copied, altered or removed without authorisation during processing, use and after recording (access control),

4. to ensure that personal data cannot be read, copied, altered or removed without authorisation during electronic transfer or transport or while being recorded onto data storage media, and that it is possible to ascertain and check which bodies are to transfer personal data using data transmission facilities (disclosure control),

5. to ensure that it is possible after the fact to check and ascertain whether personal data have been entered into, altered or removed from data processing systems and if so, by whom (input control),

6. to ensure that personal data processed on behalf of others are processed strictly in compliance with the controller's instructions (job control),

7. to ensure that personal data are protected against accidental destruction or loss (availability control),

8. to ensure that data collected for different purposes can be processed separately.

3.3.2.4.4. French Data Protection Act

The key data protection legislation is Act No. 78-17 on Information Technology, Data Files and Civil Liberties 1978 (French Data Protection and Freedoms Act) as amended and Decree No.2005-1309 implementing the French DPA. The DPA transposes into French law the requirements of the DPD and requires that organisations implementing data processing or holding data files guarantee their security. Data security should be understood as all *"useful precautions, with regard to the nature of the data and the risks of the processing, to preserve the security of the data and, in particular, prevent their alteration and damage, or access by unauthorised third parties"* (Article 34). Specific standards are not stipulated by law or binding guidance.

3.3.2.4.5. Spanish Data Protection Act

In 1999, DPD was transposed into Spanish legislation through the Data Protection Act (Ley Orgánica 15/1999 de Protección de Datos de Carácter Personal (LOPD), which establishes a general obligation on data controllers and, where required, data processors to adopt the technical and organisational measures necessary to guarantee the security of the personal data they process. Similar to the aforementioned Directives and Data Protection Acts implementing the DPD, LOPD was articulated in a technology neutral way repeating the wording of DPD without setting any specific data security standards. The need for more legal certainty in the data protection regime, particularly on issues that over the years had proven to be in need of further regulatory implementation, led the Spanish legislator to enact an ancillary legislation to LOPD, the Data Protection Regulation (Real Decreto 1720/2007 (RLOPD). Recognising the importance of the duty of security and based on the difficulties faced by data controllers, which revealed the weak points of previous data protection laws, the Regulation is particularly rigorous in the assignment of the levels of personal data security. Taking into consideration the various forms of material and personal organisation of security in common practice, it is precise in regulating the obligations associated with data processing and data files thus providing data controllers with a clear framework in which to

act. RLOPD sets out detailed and specific requirements regarding the minimum measures to be adopted in each case establishing a layered approach whereby data controllers would implement different levels of security accumulatively depending on the nature of the information they process. The Regulation classifies the security measures into three levels – basic, medium and high – and makes distinction between automated and non-automated data files.

Part VIII of the RLOPD covers in detail the security measures applicable to automated and non-automated filing systems and processing of personal data in the last two chapters, while the second chapter is dedicated to the 'security document' analysed below. The first chapter defines the level of security required in each case depending on the type of personal data processed and mandates that basic-level security measures should be applied in all the cases of data processing. In addition to basic measures, medium-level measures must be implemented when the data processed are related to financial services, public security or public tax matters, or which may allow data controllers to identify a data subject. Finally, a high level of security should be applied to databases containing sensitive information (ideology, religion, trade union membership, racial origin, health or sex life), data arising from acts of gender-based violence, and data collected for security forces without the consent of the data subjects. High-level security measures must be applied in combination with basic and medium-level measures. Article 81(4) specifically deals with two certain categories of data – location and traffic data – when processed by providers of publicly available communications services and requires the following information to be stored in each attempt to access files containing such data; identification of the user, the date and time when access was attempted, the time of access and whether it has been authorised or denied. Chapter II introduces the concept of security document as a means for protecting personal data obliging data controllers to draw up a document including the necessary technical and organisational measures imposed on the personnel with access to the information systems; controller should regularly check whether the obligations contained in the document are sufficiently fulfilled. The security measures that need to be adopted are stipulated in Articles 89-114 of RLOPD, a brief summary of which is provided below. It is noteworthy that these measures overlap with the measures described in various guidelines on how to implement data protection laws, but the main difference is that in the case of RLOPD they are binding on those to whom it is addressed and thus directly applicable and enforceable.

1. Security measures applicable to automated filing systems and data processing

i. Basic-level Security Measures

- *Obligations related to staff*: Staff obligations should be clearly defined and documented in the security document. Data controller should be in charge of staff members' security training that would make them aware of the security regulations that affect their functions and of the repercussions that incompliance might entail.

- *Incident Record Management*: A procedure for notification and management of incidents affecting personal data should be put in place and a register should be established for recording details regarding the incident (type, time etc.).

- *Asset Access Control:* Data controller should establish mechanisms so that access is permitted only to the resources required for one's functions.

- *Support Management Systems:* The documents containing personal data should allow the identification of the information included and should be accessible only to authorised personnel. The transfer of documents containing personal data, including those comprising and/or attached to emails, should be authorised by the data controller.

- *Identification and Authentication:* The data controller should implement the necessary measures that guarantee the correct identification and authentication of the users. In the case that the authentication system is based on passwords, certain guarantees should be in place to protect their integrity and confidentiality, such as encrypted password storage and mandatory password change (minimum once a year).

- *Backup Copies and Recovery:* Data controllers should establish protocols for making backup copies as well as procedures for the recovery of data that could guarantee their reconstruction to the original state in case of a loss or deconstruction. The implementation of these procedures should be monitored every six months.

ii. Medium-level security measures

The measures of this category involve all the basic security measures described above plus the appointment of a data protection officer and the execution of privacy audits every two years. Security officers are in charge of monitoring the implementation of the measures, analyse the audit reports and inform the data controller of possible measures that need to be adopted; the designation of a security officer does not entail controller's exemption of liability. At least every two years the data processing systems and data storage installations should be subject to an internal or external audit aiming to identify deficiencies in the adaptation of the measures and propose the necessary complementary or corrective measures.

iii. High-level security measures

One of the measures to be adopted is associated with the access to personal data; strict physical access control is required in the sense that premises where computers used for data treatment are located must be subject to strict physical boundaries and control checks. Moreover, the distribution of files containing personal data or the transfer of personal data through public or wireless electronic communications network should guarantee that the data are not intelligible either by encoding them or using any other mechanism for this purpose. When portal devices containing personal data are outside the installations of the data controller, the processing of such data should be prohibited in the case of devices that do not permit encoding. Another high-level security measure is the full access logging for each attempt to access personal data, which involves the storage of information such as the filing system accessed, the data and time it occurred, the identification of the user etc.

2. Security measures applicable to non-automated filing systems and data processing

The respective measures described above should also be adopted on each level of security when the filing systems and the data processing are not automated. The only measure added to guarantee high-level security is related to the storage of personal data; the cupboards and filing cabinets used to store non-automated files with personal data should be placed in areas to which access is protected by entrance doors with locks or other equivalent devices.

3.4. The example of Israel's law

Similar to the Spanish data protection law are the Protection of Privacy Regulations ("the Regulations") passed by the Israeli Parliament (the Knesset) on March 21, 2017, which impose mandatory data security and breach notification requirements on anyone who owns, manages or maintains a database containing personal data in Israel. The Regulations, which will enter into force in late March 2018, classify databases into four categories, each subject to an escalating degree of information security requirement. The regulatory model is structured as follows:

i. Security requirements applicable to databases held by individuals

Databases pertaining to this category are either maintained by an individual, in the sense of a sole proprietor, or are held by a corporation with a single shareholder, with no more than three persons having access credentials. In addition, the following three conditions need to be met for a database to be treated as such: the number of data subjects contained within a database should not exceed 100,000; the data must not be subject to professional confidentiality obligations under law or codes of ethics (e.g. a database maintained by a sole-practitioner attorney); databases must not be used to make information available to third parties (e.g. databases used to provide direct marketing services). Owing to the low risk involved in the processing of such databases, the security requirements provided for this category are lenient, including physical and communication security; access credentials, authentication and user administration; restrictions on using portable devices; segregation of systems; drafting a database specification document; and documenting information security incidents.

ii. Security requirements applicable to all databases except those held by individuals

The requirements imposed by the Regulations apply irrespective of the security level to which databases are subject to and stipulate certain security measures as described below:

- *Drafting a database specification document:* Database owners are required to draft a document containing details relating to the data collected, the purposes of the processing activities, the types of data processed, the identity of the database manager, its data security officer and its holders (i.e. those who usually possess a

copy of the database and are entitled to use it), as well as the processing activities carried out by other processors, the security risks involved and possible ways of mitigating them.

- *Physical security:* The computer systems of a database must be kept in a secure place to which only authorised individuals can have access.

- *Data security officer:* Companies, financial institutions and public agencies maintaining five or more databases are required to appoint a data security officer, who is in charge of establishing data security protocols, preparing an ongoing plan to review compliance with the Regulations and presenting the outcome of the review to the data manager and its supervisor.

- *Data security protocols:* Security protocols should be established containing instructions with respect to the physical and environmental security of the database's premises, the administration and use of portable devices, the access credentials, the measures that need to be implemented in order to safeguard the database's computer systems, and the risks to which the database is exposed. Additionally, the protocols should also include a security incident response plan, which should be layered varying in accordance with the incident's severity and the degree of sensitivity of the database.

- *Mapping the database's computer system:* Database owners should compile an updated list of devices and components that comprise the database's computer system, which should also include hardware and software components and describe the architecture of the system in which the database is installed.

- *Access credentials, authentication and user administration:* Database owners should take measures to ensure that access credentials are assigned in accordance with each user's duties and maintain a list of those with access credentials.

- *Documenting information security incidents:* Database owners should maintain documentation for any security incident likely to result in a data breach.

- *Portable devices:* Restrictions must be placed on the use of portable devices with database-related computer systems (e.g. laptops, smartphones, memory sticks). Such restrictions should be placed in accordance with the database's degree of sensitivity.

- *Segregation of systems:* Database owners must segregate database-related computer systems from other computer systems to the extent possible.

- *Communication security:* When database-related computer systems are connected to the Internet, appropriate security measures must be implemented to safeguard against unauthorised access and malware. Examples of such measures include encryption of personal data when transmitted over the Internet and authentication of employees in case of remote access to the database.

- *Outsourcing:* When an outsourced data processing provider is engaged, due-diligence review of the risks of such an engagement must be conducted in advance. The contractual engagement should address issues such as the type of the data processing to be performed, the purposes for which the data will be used, the period of engagement and return of the data upon conclusion of the engagement.

iii. Security requirements applicable to databases subject to the basic level of data security

This category of databases consists of all the databases that do not fall within any of the other categories. In addition to the abovementioned requirements applicable to all databases (except those held by individuals), the Regulations introduce the following requirements for this category:

- *Annual reviews:* Database owners must review the security protocols annually to determine whether any updates are needed.

- *Training:* Database owners should provide users with access credentials with training with respect to the security protocols and security requirements before granting access privileges or in case of altering their scope.

- *Record keeping:* Information relating to compliance with the Regulations should be retained for 24 months. Examples of such information include documentation of security incidents, access privileges and authentication measures, information about communication security.

iv. Security requirements applicable to databases subject to the intermediate level of data security

The types of databases described below fall within this category: databases maintained by public agencies; databases whose purposes include making information available to other parties; databases to which more than 10 people have access credentials; databases containing special categories of data. Special categories of data include, among others, medical or health information, genetic or biometric data, information about an individual's religious beliefs, faith, political opinions or criminal convictions. Special categories of data also extend to cover financial information about an individual's financial obligations, solvency or financial status, as well as information regarding a person's consumption habits that may be indicative of the aforementioned types of data. For databases subject to the intermediate level of security the following requirements apply in addition to the requirements applicable to all databases:

- *Physical and environmental security:* Any equipment brought in or taken out of the database's premises as well as access to such premises must be monitored.

- *Extended security protocols:* Security protocols must address, among other issues, access controls, authentication measures, backup procedures and periodic audits.

- *Authentication:* Users with access credentials should be authenticated with physical devices, such as smart cards. In addition, a protocol must be established as a means of identification, password or frequency change, and response to errors in access controls.

- *Monitoring access:* An automated mechanism must be established in order to monitor access to the database and the logs should be maintained for at least two years.

- *Periodic audits:* At least once every two years an audit (either internal or external) must be conducted, which should include a report assessing the compliance of the security measures with the security protocols, identifying deficiencies and proposing ways of remediating them. The report must be reviewed to determine the need for updating the database's specification document and security protocols.

- *Backup and recovery:* A backup and recovery plan must be established.

- *Security incidents:* Security incidents must be reviewed at least once a year to determine the need for updating security protocols.

- *Data breach notifications:* In case of a severe data breach, where a material part of the database was accessed or used without authorisation or the database's integrity was compromised, prompt notification must be provided to the privacy regulator, who may order the database owner to notify all affected data subjects. The Regulations do not prescribe any sanctions for violating the breach notification requirement.

v. Security requirements applicable to databases subject to the high level of data security

Databases pertaining to this category should meet the following criteria: databases containing special categories of data, in which either the number of data subjects is 100,000 or more, or to which more than 100 people have access credentials; databases whose purposes include making information available to other parties, in which either the number of data subjects is 100,000 or more, or to which more than 100 people have access credentials. In addition to the requirements applicable under the basic and intermediate level, the following security requirements apply to this category:

- *Risk assessment:* Database owners must carry out a risk assessment once every 18 months in order to identify security risks and deficiencies, which should be remedied, while security protocols should be updated accordingly. Risk assessment can also be leveraged to satisfy the requirements of periodic audits.

- *Penetration tests:* The computer systems of the database must be subjected to penetration tests once in 18 months in order to evaluate their robustness in the face of internal and external security risks.

- *Security incidents*: Security incidents must be reviewed at least once every calendar quarter and the need for updating security protocols must be assessed.

- *Data breach notification:* The breach notification requirement applies to any severe breach in which any portion – not only a material part - of the database was compromised.

3.5. Technology Neutrality vs. Privacy by Design

3.5.1. Privacy by design: Overview

The speed at which information is being moved into digital environments where automated processes are in place, coupled with the fact that the mere implementation of privacy enhancing technologies (PETs) has been proved insufficient to extensively address the challenges posed to individuals' right to privacy, calls for a reconsideration of what personal data is collected, stored, used and then protected in these novel environments (Morton & Sasse, 2012). To address these concerns, a holistic approach was introduced under the term *Privacy by Design (PbD)*,[22] which places emphasis on the need for data protection safeguards to be built into products and services from the earliest stage of their development. The Canadian Commissioner for Ontario, Dr. Ann Cavoukian, coined the term (Cavoukian, 2006), which was recognised as the global privacy standard in 2010,[23] but only recently has it received more attention in terms of its inclusion as a positive requirement into the EU (as well as the US and Canadian) data protection framework. The most salient of the seven foundational principles that need to be practiced for the objectives of PbD to be accomplished is the requirement to move from a reactive to a proactive approach (Cavoukian, 2011a). Instead of trying to mitigate the risk after a privacy-invasive event has occurred, PbD encourages organisations to prevent the risk by designing the architecture of the system in a privacy-respective manner. PbD can be seen as a valuable tool to protect data and privacy as well as build trust in the systems, since it requires organisations to proactively embed privacy features into their business models not only in the beginning of the design process but throughout the whole life cycle of the process development in order to ensure that the processes put in place remain relevant as risks to data evolve (Krebs, 2013). In this way, personal data is adequately protected during every step of the data processing activities – transmission and storage of, and accessing to, the data (Gutwirth et al, 2015).

[22]<https://www.privacybydesign.ca/>

[23]*See* Privacy by Design Resolution, 32nd International Conference of Data Protection and Privacy Commissioners, 27-29 October 2010, Jerusalem, Israel, available at https://www.ipc.on.ca/site_documents/pbd-resolution.pdf.

3.5.2. Should privacy by design become a legal obligation?

Despite the fact that PbD is often cited as a best practice or even as the 'golden standard' in data protection, and guidelines have been published on how PbD should be understood, it has not yet formed an explicit part of the legislative scheme in the EU. Whereas the existent data protection laws aim at promoting the PbD, in practice they have not been sufficient in ensuring that data protection is embedded in ICTs since the need for taking data protection into consideration at the design stage of a system has merely indirectly been addressed in the EU legal framework. Calls for the introduction of PbD into legislative frameworks have lately received more attention due to the process of reforming the EU data protection regime. However, conflicting interests have as a result not all stakeholders to share the same view on the issue at stake. The debate surrounding a legislated PbD requirement revolves around three main perspectives: some advocate the integration of PbD requirement in the legislation including certain mandatory technological features for those technologies that have the most privacy-intrusive potential; others argue that PbD should become a legal obligation only as a general principle of data protection law; others see PbD more as a self-regulatory initiative rather than as a part of legislation (Krebs, 2013). To start with, Article 29 Working Party has opined that PbD should be a legislative requirement not only as a general principle but also as a binding requirement imposed on all the actors concerned – data controllers, data processors, designers and purchasers of systems or applications (Article 29 Data Protection WP, 2009). The view that regulators should take a clear position on the importance of PbD by incorporating it in legislation seems to be gaining ground; PbD should be explicitly mentioned in privacy laws as a general organisational principle, while specific technological requirements should be provided when the nature of the data and the systems calls for them. In a recently published document, the European Network and Information Security Agency (ENISA) expressed the same opinion highlighting that PbD principles and mechanisms can be effective only if reflected in legislation. Likewise, recognising PbD as 'an essential component of fundamental privacy protection', some Data Protection Authorities (DPAs) contend that PbD is a concept that needs to be encouraged through legislation; differences in the notion of privacy across cultures, though, might result in divergent DPAs' opinions. As expected, the strongest objections have been raised by industry players, who fear that prescriptive technological mandates would impose extra technical and economic burdens

with detrimental effects on competition and innovation.[24] As explained by Microsoft in a Consultation to the European Commission relating to data protection approaches, 'an industry-wide obligation' should be applicable to the ICT industry; PbD obligations are welcomed as long as they do not take the form of design mandates or technology preferences.[25] Closely aligned with this view has been the UK Information Commissioner's opinion, according to which high-level regulation and self-regulation initiatives should be preferred over technologic-specific regulation.[26]

3.5.3. Proposal for a General Data Protection Regulation

The need for increased legal certainty in the EU context has resulted in the drafting of a seemingly technology specific legislation, the proposed EU General Data Protection Regulation, whose aim is, among others, to provide a high degree of legal certainty for both data controllers and data subjects. As stated by the European Commission, the current legal framework on the fundamental right to personal data protection has not managed to prevent legal uncertainty, while it has been highly criticised by economic stakeholders who seek for more certainty in terms of their legal obligations.[27] Although the principles and objectives as outlined in the current framework remain valid, the rapid development of new technologies – especially online – calls for a new legislation able to better respond to those challenges. The need for specific legal norms stems from the effects of the personal data processing systems (PDPS), which often violate the substance of the right to data protection[28]. As described below, the GDPR is viewed as a compensation for the detrimental impact of the usage of

[24]*See* e.g. Vodafone, Vodafone's Response to the Consultation on the Commission Communication on *'A comprehensive approach on personal data protection in the European Union'*[COM (2010) 609/3], European Commission, available at http://ec.europa. eu/justice/news/consulting_public/news_consulting_0006_ en.htm; Facebook, Submission to the Consultation on the Commission Communication on *'A comprehensive approach on personal data protection in the European Union'* [COM (2010) 609/3], European Commission, available at http://ec.europa.eu/justice/news/ consulting_public/news_consulting_0006_en.htm.

[25]Microsoft, Submission to the Consultation on the Commission Communication on *'A comprehensive approach on personal data protection in the European Union'* [COM (2010) 609/3], European Commission, available at http:// ec.europa.eu/justice/news/consulting_public/news_ consulting_0006_en.htm.

[26]*See* Response to the Minstry of Justice's Call for Evidence on the Current Data Protection Legislative Framework, December 2010, avaialable at https://www.cfoi.org.uk/pdf/DPAresponses.pdf.

[27] Resolution of the European Parliament on the on the Communication from the Commission to the European Parliament and the Council – An area of freedom, security and justice serving the citizen – Stockholm programme, adopted 25 November 2009 (P7_TA(2009)0090), available at http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2009-0090+0+DOC+XML+V0//EN.

such technologies. In addition, the direct applicability of the Regulation provides an extra guarantee for greater legal certainty.

One of the most salient measures explicitly introduced for the first time in the EU legal framework in order to achieve the aforementioned goal is the legal obligation of *Data Protection by Design* (DPbD). The GDPR seems to use the terms 'privacy by design' and 'data protection by design' synonymously (ENISA, 2014a). DPbD can be seen as a measure of legal certainty because under this obligation the requirement of data protection is phrased in technical standards. At first sight, DPbD seems to infringe upon the law-making principle of technology neutrality as it does not merely deals with the implementation of technologies but it also interferes with technology design. As a matter of fact, however, the general wording of the respective provision contains the element of technology neutrality as it abstracts from specific technical solutions. Without imposing any particular technologies, Article 25(1) of the GDPR requires data controllers to implement appropriate technical and organisational measures and procedures, both at the time of the determination of the means for processing and at the time of the processing itself, "in such a way that the processing will meet the requirements of this Regulation and ensure the protection of the rights of the data subject." The obligation to implement DPbD does not introduce technology-specific rules, in the strict sense, as it neither stipulates which specific technologies it addresses nor explains which particular technologies should be deployed to achieve DPbD.[29] Instead, the third and fourth paragraph of the same article empower the Commission to specify further criteria or requirements with respect to the appropriate measures or lay down technical standards that meet the requirements set out in the Regulation. Article 25 not only does not seem to violate the principle of technology neutrality but, on the contrary, it is in accordance with one of its rationales – the rationale of sustainability. The Commission can indubitably react faster than the EU legislator in terms of coming into an agreement with stakeholders on the technical and organisational standards to be developed thus ensuring high-speed adaptability to high-speed developments in the ICT area.[30]

[29] See the debate on explicitly listing pseudonymisation of personal data as a key option of implementation in Article 23, Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) [First reading], 13772/14, Brussels, 3 October 2014, available at http://register.consilium.europa.eu/doc/srv?l=EN&f=ST%2013772%202014%20INIT.

[30] It can be argued that extended powers have been assigned to the Commission as it is allowed to intervene as a legislator through delegated acts which have supremacy over national laws and constitutions.

Another principle introduced by the GDPR is *Privacy by Default*, which is seen as the ideal condition for the goals of *Privacy by Design* to be achieved. Article 25 (2) requires data controllers to provide data subjects with the highest level of data protection by default, while it is at the data subject's discretion to erode the level of protection by explicitly opting out of particular safeguards. Under the principle of Privacy by Default, once a customer acquires a new product or service, the strictest privacy settings should automatically apply and therefore no manual change should be required on the part of the user (Hansen, 2013). Privacy by default is based on the insight that the user must already be protected against privacy risks in the default settings. As provided in Article 25 (2), only personal data necessary for a specific purpose should be processed and by default only be kept for the amount of time necessary to provide the product or service. It is noteworthy that the provision specifically identifies and prohibits services that by default make personal information accessible to an indefinite number of individuals. The Privacy by Default principle is likely to be highly beneficial to data subjects, particularly in the area of social networking platforms, where more personal information is disclosed to the public than necessary to provide the user with the service, while it is likely that personal data is stored for a long period of time and easily accessed by unauthorised individuals (Danagher, 2012).

3.5.3.1. Data controllers' discretion

Based on the premise that compliance obligations imposed on data controllers should be proportional to the specific processing activities, the Council of the EU has opted for a risk-based approach to compliance. The choice of the word 'appropriate' when speaking of the measures that need to be taken in the Articles 25(1) and 32(1) of the GDPR is indicative of the Council's intention to allow data controllers to exercise discretion and flexibility in assessing how to address their compliance responsibilities in the context of their particular business. In particular, with respect to the obligation of data protection by design, the wording of Article 25(1) clearly shows that a broad view of 'design' has been preferred, which does not only mean the implementation of PETs; more than that, it involves the integration of technical and organisational measures into the business models of data controllers (Hildebrandt & Tielemans, 2013). The requirements of data protection and security by design seem to have been more adaptable to the context of data controller's business model and therefore the level of measures considered appropriate is up to the controller to decide. The extent of the data controller's obligations will depend, among

others, on the state of the art and the costs of implementation; in case that the highest level of data protection and security measures is technically and economically infeasible, then the data controller cannot be held liable (Hildebrandt, 2013). In addition, it will be open to the controller to define the purpose of the processing activities and whether it is necessary to process, collect and store the data for that purpose. As expected, objections have been raised regarding the risk-based approach that the Council intends to adopt. In a debate over the progress of the GDPR,[31] fears were expressed that this approach might weaken the EU's ability to provide strong data protection rights for individuals, since the great discretion granted to data controllers might result in a wide margin of interpretation, rendering it questionable whether the necessary data protection guarantees can be provided.

3.5.3.2. Targeting only data controllers and not technology developers

It is noteworthy that the obligation to implement DPbD targets only the users of data processing technologies, that is, the data controllers, whereas no burden of responsibility is assigned to the designers or manufacturers of such technologies. The provision of the GDPR does discriminate between technology developers who are also data controllers and those who are merely technology vendors. This option can be viewed as an indirect means for making data controllers force developers to come up with the right types of technologies by motivating them to invest in compliant design (Hildebrandt & Tielemans, 2013). However, an inconsistency should be noticed at this point; the obligation to design technologies with data protection features embedded applies to a company which also carries out data processing activities, but there is no such obligation on technology developers when a company only invests in technologies built by others. It has been argued that technology vendors should also be liable for defects in their products that result in their users' security exposure and therefore the same obligation must also be attributed to technology developers, even if they are not data controllers (Klitou, 2014). This argument is reinforced by the reluctance observed on the side of developers to take into account the individuals' right to data protection due to the lack of economic incentives. The lack of liability allows those who fabricate and sell data processing systems to produce more products faster with less concern about safety and data protection (Schneier, 2003). Imposing the same obligation to

[31] *See* Can the next EU Regulation guarantee data protection for all?, January 20, 2015, available at http://www.vieuws.eu/live-panel-debate/debate-can-the-next-eu-regulation-guarantee-data-protection-for-all/.

technology manufacturers and holding them liable for security vulnerabilities would act as the market force that would incentivise them to invest in secure development processes (Schneier, 2008). Besides, responsibility should be assessed wherever it will most effectively reduce the risks inherent in defective products. As expected, there is strong opposition to the idea of liability on behalf of security software designers, manufacturers and sellers who argue that, first, it is impossible to foresee the normative impact of a technology not yet operational and, second, that those who benefit from the use of data processing technologies should also bear the cost of responsibility when infringing data protection legislation (Ryan, 2003).

3.6. Conclusion

When assessing the adequacy of the EU data security legislation, there is often a discussion about whether EU laws should be technology neutral or they should mandate a specific list of security measures to be implemented by data controllers. The security paradox lies in the fact that in security much depends on the technical details of implementation, but these details are hard to capture in high-level legislation. On the one hand, one could argue that the organisations and technologies involved are too diverse to allow for a single checklist for the entire sector. The rapid change of ICTs, coupled with the rapidly evolving capabilities of attackers, call for sustainable laws and flexible high-level regulation that can be easily adaptable to technological advances and applied to a wider number of organisations. On the other hand, high-level standards inevitably leave a lot of significant technical details unaddressed thus failing to create clarity about what measures data controllers need to take to be compliant. Such is the case of the current EU legal framework on the security obligations incumbent upon data controllers, as it has not managed to provide sufficient certainty regarding their obligations. As mentioned above, the relevant Directives are written in a technology neutral way allowing data controllers to determine what security measures would be ‘appropriate’ depending on their particular business. In the context of the current legislation, the obligation to implement the appropriate technical and organisational measures can be interpreted as a legal obligation to act reasonably. In other words, organisations are required to reflect on the technologies in the marketplace and make a conscious decision in terms of what technologies they could reasonably apply to safeguard their databases and applications.

However, the great discretion granted to data controllers in assessing how to address their compliance responsibilities is likely to result in a wide margin of interpretation. In an attempt to encourage competition in electronic communication markets, the EU regulator appears to attach more importance to the interests of market players to the detriment of personal data protection. Given that nowadays there is a new model of governance, according to which all fundamental rights go through private infrastructure, while private organisations seem to rely on their privacy policies rather than on the law, the extent to which power should be left to companies empowered by processing of personal data should be carefully assessed. The security best practices formulated by organisations such as Internet Service Providers' Associations are not compulsory for their members, while the same applies to the non-binding recommendations issued by European bodies. The need for legal specificity as to data security obligations is reinforced by the implementation of invasive security techniques, such as deep packet inspection, that often infringe upon the right to data protection. The lack of clarity of the legal framework concerning the admissibility of Internet traffic inspection methods, combined with the high degree of flexibility granted to data controllers, render it questionable whether the necessary data protection guarantees are actually provided. As analysed above, when technology designs threaten fundamental human rights, specific provisions might be necessary to retain the substance of the legal right they support. Combining technology neutral legislation with certain more detailed provisions could be the ideal way of dealing with the legal security obligations of data controllers. Such is the case of the Spanish Data Protection Act, which lays down these obligations by specifying the assets that need to be protected and the means of safeguarding the data they process without mandating particular technical features of the technologies used. As a result, law remains open to new technological designs while data controllers are obliged to implement certain measures in order to achieve the goals set out by the Act.

In an attempt to force data controllers to bolster their data processing security practices, the European Commission has proposed a General Data Protection Regulation as a compensation for the detrimental impact of the usage of privacy-invasive technologies involved in the processing of personal data. However, the GDPR seems to fail to provide the necessary legal certainty with respect to the security measures that data controllers are required to adopt since the wording of Article 25(1) also adopts a broad approach allowing data controllers to determine which security techniques are considered appropriate in the context of their business. The main difference the GDPR might bring in the data protection regime is the

explicit requirement of data protection by design (DPbD), which can be viewed as a guarantee that this principle will be embedded in technologies from conception to finalisation. In order to establish data protection by design as an organisation's default mode of operation, not merely as an organisational best practice, data protection laws need to include a process through which a product or a system could be prevented from entering the market until it is sufficiently data protective. There have been concerns, however, that making DPbD an explicit part of legislation would likely not be enough to ensure the desired level of data protection unless this requirement is also addressed at software manufacturers. Despite the fact that several guidelines have been published on how DPbD could be understood, many companies design systems or products that do not have adequate protection in their architecture either because they are not familiar with technologies encompassing data protection principles or because compliance with DPbD could potentially become complex, resource-intensive and expensive. System developers mostly focus on realising functional requirements and, as a result, they hardly consider other demands, such as privacy or security guarantees. Another concern with regard to the effectiveness of DPbD in the GDPR is linked to the limited means that data protection authorities currently have at their disposal to assess the degree of implementation of data protection principles in ICT systems. The conceptual difficulties in guaranteeing data protection properties in dynamic systems, that is, systems that adapt to changing requirements, add a further burden in the process of evaluating compliance with DPbD. Therefore, in order for DPbD requirement to make a significant change in the current EU legal framework, it is primarily important to reach a common understanding of what DPbD entails. Moreover, every actor in the life cycle of a technology should be accountable from a data protection perspective, while it is the regulator's mandate to ensure that adherence to the legislative requirements can be traced and actually enforced.

CHAPTER 4: Deconstructing the Technological Neutrality of the Data Protection Directive

4.1. Introduction

The aim of this chapter is to provide insights on the mindset underlying the EU data protection legal framework and thus answer the question of whether the DPD meets the conditions required for a legal approach to be considered as technology neutral. To this end, the fundamental block of the data protection regime is studied, that is, the concept of personal data, which determines the scope and boundaries of the relevant laws. The identifiability notion, on which the definition of personal data is based, and the distinction between personal and anonymous data, which implies the robust anonymisation assumption, reveal the digital technological paradigm into which the concept of personal data is rooted. In order to appraise the effectiveness of the anonymisation techniques in irreversibly converting personal data to de-identified data, infamous re-identification cases are briefly presented. The challenges posed by the current state of technological advancement in respect of the flaws of the underlying robust anonymisation assumption raise the question of whether the concept of personal data needs to be abandoned or reconsidered in a context where anonymisation will not serve anymore as a safe harbour from the entire application of data protection rules. The chapter proceeds with the analysis of the second crucial definition – that of the personal data processing – and demonstrates the linear consequence in which data processing activities have been organised. In spite of the seemingly technology neutral character of the definition, which can, in fact, easily accommodate future technologies, carefully reading the list of the possible operations performed upon personal data can reveal the underlying technological assumptions of the definition. Another point made in relation to this definition is the fact that the law relies data processing upon the 'informed, explicit and unambiguous' consent of the data subject, thus failing to take into account the currently deployed industry's practices that are indicative of the absence of the user's physical ability to exercise control over their data processing on the ground.

4.2. Technological Assumptions Hidden in the Concept of Personal Data

4.2.1. The digital mindset underlying the identification-based definition of personal data

The fundamental building block of the entire EU data protection regime is the precondition that the data being processed must be personal. Pursuant to Article 2(a) of the DPD, personal data is defined as "any information relating to an identified or identifiable natural person ('data subject')". The definition of personal data, which is bundled with the definition of the data subject, indicates that the criterion applied by the DPD is the identification of the subject rather than the content of the data. Contrary to the sectoral model adopted by the U.S., according to which privacy protection takes place in an enumerated set of cases[32], each of which is protected by a specific federal law, the EU omnibus approach, in its choice of identification as the basic criterion for its regulatory scheme, aims at protecting personal data across the board, not only in a specific sector or context. On top of the identification-based layer, there is a second content-based layer which requires additional legal attention to the processing of specific categories of data, namely "data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life" (Article 8 (1)). Accordingly, the U.S. approach does not end with the content-based criterion since there are informational privacy laws that apply the European standard of identification.

The reference to identification, which involves "describing a person in such a way that he is distinguishable from all other persons and recognisable as an individual" (European Union Agency for Fundamental Rights, 2013), includes not only the cases where the identity of an individual is already distinguishable ('identified data subject') but also the cases where a person "can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity" ('identifiable data subject') (Article (2)(a)). Therefore, for the DPD to apply, it is not required that the individual be identified only at the level of his name; identification based on the above factors would also suffice. 'Identifiability' implies the

[32]*See*, for example, Health Insurance Portability and Accountability Act of 1996 (HIPAA), Pub. L. No. 104-191, 110 Stat. 1936 (1996) (health-related data); Financial Services Modernization Act of 1999, Pub. L. No. 106-102, 113 Stat. 1338 (1999) (known as Gramm-Leach-Bliley Act) (financial data); Video Privacy Protection Act, 18 U.S.C. § 2710 (2006) (data about video rentals.); Family Educational Rights and Privacy Act of 1974 (FERPA), 20 U.S.C. § 1232g (2006) (data about students' records or information).

possibility for identification to occur by associating, for instance, certain information being processed with a particular person (Esayas, 2015). In this regard, the DPD's approach is considered to be more advanced inasmuch as it acknowledges that seemingly innocent pieces of data can be combined to form a whole that is greater than the sum of its parts (Birnhack, 2013). The digital mindset reflected in the European approach, as opposed to the analogue mindset of the U.S. approach, refers to the creation of individual profiles as the result of joining separate bits together, which are further analysed. Digital technologies enable the aggregation and analysis of mass quantities of personal data in a way very different, in nature and quantity, than the equivalent analogue activities, in that data does not need to be structured to be subject to data mining (Solove, 2004).

Another argument in favour of the fact that the definition of personal data is rooted within a digital technological paradigm, which lies at the heart of this definition, is that data rendered anonymous is not protected by the principles of the DPD as long as the data subject is no longer identifiable. Recital 26 contains a major technological assumption, on which most data protection laws have been based, that complete anonymisation is possible to achieve. Provided that anonymisation technologies can irreversibly prevent identification, anonymisation can serve as a safe harbour from the entire application of data protection rules. Ohm (2010) argues that the choice of the EU lawmakers to rely upon anonymisation as 'a silver bullet solution' should be seen as an attempt to strike a balance through the power of technology. In order to combine both the benefits of information flow and strong assurances of privacy, legislators chose anonymisation to avoid engaging in overt balancing between countervailing values such as security, innovation and the free flow of information. Illustrative of this attempt is the long-lasting debate in the internet privacy context between the EU and companies like Google, Yahoo and Microsoft over the security measures that should be implemented to protect databases used to track online users' behaviour. The focal point of the debate was the protection of stored IP addresses. While all the above companies used anonymisation to protect users' privacy, Microsoft and Yahoo were throwing away the entire IP addresses, unlike Google that stored the three first *octets* – equal pieces which an IP address is composed of – and deleted the last piece for more effective protection as claimed (Comer, 2006). In essence, the debate over users' privacy involves cost-benefit balancing arguments as it turns out to be a debate over balancing the benefits stemming from the innovative techniques of analysing online behaviour, on the one hand, and the possible harm caused to the individuals concerned by the disclosure of such information (Ohm, 2010).

Irrespective of whether the drafters of the DPD could predict the digital environment or their reasoning for choosing the identification-based definition was to avoid the difficulty of reaching an agreement on which categories of data were worthy of legal protection, the DPD is informed by a digital concept. Even though the definition of personal data is technology neutral with respect to the language and references used, there is a central hidden technological assumption in the DPD, which assumes a technology of a particular capability, namely irreversible anonymisation. The danger involved in such an assumption is that the definition should be understood within a certain technological paradigm and hence it is sufficient as long as the paradigm remains the same. Once, however, the paradigm changes, the concept of non-identification will collapse and lawmakers will need to find new ways to regain the lost balance. As demonstrated below, the current state of technological advancement renders the notion of irreversible anonymisation largely obsolete thus revealing the flaws of the underlying assumption on which the EU data protection regime has been founded.

4.2.2. Personal data as a troubled concept for framing data protection regulation

4.2.2.1. The robust anonymisation assumption

The development of computerised record systems and techniques of digital data analysis that introduced new ways to link data to individuals has resulted in the emergence of the concept of personal data as a concept pervasive in both legal and technological discourse on data protection. Personal data is the most central concept in the EU data protection legal framework as it defines the scope and the boundaries of the relevant laws whose application is triggered only when the data involved in the processing activities are identifiable (Schwartz & Solove, 2011). Based on the assumption that in the absence of personal data there is no actual harm, data protection laws leave the collection, use and disclosure of non-personal data unregulated. Behind the concept of personal data is hidden the dominant role that anonymisation plays in contemporary data protection practices as the most effective technique for converting personal data into non- personal data. The techniques employed in data anonymisation are designed to turn data into a form which does not identify individuals and where identification is not likely to take place.

The faith in robust anonymisation, which has thoroughly infiltrated the EU data protection system, stems from the assumption embraced for decades by both technologists and

regulators that anonymisation can guarantee complete protection of personal data. The robust anonymisation assumption has worked as a data protection panacea for various parties. Legislators could balance data protection with other interests, such as the free flow of information, by deregulating the processing of anonymised datasets, data administrators could safely share data with third parties and data subjects could rest assured that their data would remain secret. The fact that most discussions about anonymisation often merely focus on the technical details, with which the majority of individuals is not familiar, gives a false sense of security to them as they are not capable of fully comprehending the potential uses of anonymised data and the dangers involved and thus they tend to downplay them (Oswald, 2014). One of such potential dangers is related to the re-identification of anonymised data. The emergence of the re-identification science, which involves the development of data linkage techniques that manipulate databases to determine the identity of individuals, undermines the faith placed in robust anonymisation and disrupts the data protection landscape.

In its Opinion 5/2014, the Article 29 WP suggests that the underlying rationale of Recital 26 of the DPD is that the outcome of data anonymisation ought to be "as permanent as erasure" for the data not to be subject to the DPD's provisions. In assessing the robustness of each anonymisation technique, the following three questions should be taken into consideration: whether it is still possible to single out an individual; whether it is still possible to link records relating to an individual; whether information regarding an individual can be inferred. Applying the above questions to the anonymisation techniques currently employed, the WP concludes that some techniques show inherent limitations, while all of them fail with certainty to meet the criteria of effective anonymisation (although the technique of differential privacy[33] could have satisfying results). Moreover, according to the aforementioned Opinion, data should be treated as not truly anonymised and hence personal data in case the data controller shares part of the dataset (by removing or masking the identifiable data) to another party without having first deleted the identifiable data. However, this approach overlooks the likelihood that the original data may not assist in associating the anonymised data to a certain individual. Different is the approach adopted by the UK ICO

[33]*Differential privacy* has been described as "a set of techniques based on a mathematical definition of privacy and information leakage from operations on a data set by the introduction of non-deterministic noise. Differential privacy holds that the results of a data analysis should be roughly the same before and after the addition or removal of a single data record (which is usually taken to be the data from a single individual). In its basic form differential privacy is applied to online query systems, but differential privacy can also be used to produce machine-learning statistical classifiers and synthetic data sets" (Garfinkel, 2015).

(2012) which does not consider the disclosure of anonymised data as a disclosure of personal data even in the case where the data controllers hold the key to enable re-identification. As the following analysis will demonstrate, achieving irreversible anonymisation of the data that would not permit re-identification is considered very difficult from a technical point of view. Thus, if the acceptable risk threshold is zero, there seems to be no existing technique that can achieve the required degree of anonymisation. Legal perspectives on anonymisation are often solely concerned with the question of whether a certain technique meets the criteria for providing a safe harbour from the entire application of data protection rules without taking into consideration the risks involved in the implementation of such measures (Esayas, 2015).

This 'all-or-nothing' approach, according to which data is either anonymised or not, appears to contradict with the concept of negligibility prevailing in engineering that suggests that in practice there is no zero-risk situation, and therefore data should be considered anonymised in case of a low identification risk (ENISA, 2012). It could be argued that the latter approach is also reflected in Recital 26 of the DPD, according to which "all the means *likely reasonably* to be used either by the controller or by any other person to identify the said person" should be taken into account in order to determine whether a person is identifiable.The term *'likely'* refers to the probability of identification, while the term *'reasonably'* represents the difficulty of identification, for example, in terms of costs, time required and available technology (Bygrave, 2002). The fact that the terminology used by legislators is common in the field of the risk management reasoning indicates that DPD itself does not require a zero-risk approach but rather recognises the potential for a certain level of acceptable risk of identification (Emam & Alvarez, 2014). In other words, data anonymised to the effect that it is extremely difficult to associate to an individual could be processed without any need for compliance to the legal requirements of the DPD. Therefore, information should be regarded personal data only in the case of a "sufficiently realistic" risk of identification (Hon et al., 2014). On the contrary, when, given the technology, the costs and the time required to associate the data to a particular individual, the risk is "remote or highly theoretical" then data should not be treated as personal. The same approach has been adopted by the ICO which holds the view that anonymised data is not required to be completely risk free (ICO, 2012). It should be noted, however, that it is crucial that the application of anonymisation techniques has been engineered appropriately for the data to be processed without the need to adhere to the legal requirements of the DPD. Otherwise, any doubts regarding the effectiveness of anonymisation techniques should be interpreted as involving the processing

of personal data. Such is the case of France, for instance, where the legislators seem to ignore the 'likely reasonably' test and choose the data to be treated as personal even if it is extremely difficult to re-identify the data subject or re-identification is unlikely to take place (Esayas, 2015).

Nonetheless, even if irreversible anonymisation does not constitute a prerequisite for data to be considered non-personal and thus not subject to the DPD's provisions, the developments in powerful data analysis techniques increase the identification potential. As mentioned above, identifiability does not only refer to the possibility of retrieving one's name but also involves the potential for singling out an individual, linking two records within a dataset (or between two separate datasets) and inferring information from such dataset. Therefore, simply removing directly identifying elements would not suffice to make the identification of an individual no longer possible. Technological advances reveal the "inherent residual risk of re-identification", which is still present even when it is no longer possible to retrieve an individual's record, due to the ability to glean information about that individual with the help of other sources (Article 29 WP, 2014). Residual risks include, among others, the combination of an anonymised portion of a dataset with a non-anonymised portion or the possible correlations between attributes (e.g. between health records and geographical location). As a result, even effectively anonymised datasets can be combined with another dataset and lead to an individual's re-identification. The dynamic nature of identification risks, which increase over time along with the development of re-identification techniques, as well as the ubiquity of available information indicate the risk-based nature of anonymisation as it depends on a variety of factors that are difficult to quantify, such as the re-identification technologies and the amount of information that will be available in the future. Considering what has become in the current state of technology 'likely reasonably' and given the increasing low-cost availability of technical means to identify individuals in datasets as well as the increasing availability of datasets (e.g. Open datasets), 'likely reasonably' proves to be easily attainable (Article 29 WP, 2014).

4.2.2.2. Potential for traceability

The so-called 'anonymity myth' lies in the impression held by a large number of Internet users that there is no risk of identification as long as one does not reveal his identity while communicating online or surfing the Web. However, those users fail to realise that the

protection provided by such 'anonymity' is no more "than a veil over one's face that can readily be lifted" (Schwartz & Solove, 2011). The reason behind this misunderstanding is related to the prevailing confusion between the ability to hide one's identity for a short period of time – temporary anonymity – with the ability to permanently keep one's identity secret – actual 'untraceability'. The latter proves to be impossible due to the 'built-in identifiability' of cyberspace (Schwartz & Solove, 2011). The fact that whenever one enters cyberspace the potential for traceability exists should be attributed, amongst other factors, to the entry threshold, that is the Internet Protocol (IP) address. The identification of a seemingly anonymous user can easily follow from the IP address - a numeric identifier assigned to every device when connected to the Internet, which acts like any address in that it enables the correct delivery of data (Comer, 2006). IP addresses allow, for example, users to connect to the right Web page when typing a URL as the numerical IP address is translated to and from the alphabetical URL by the Domain Name System (DNS). The public IP address a user is allocated by his ISP may be permanent (static) or temporary (dynamic). Whereas in the case of dynamic IP addresses the device is assigned a new IP address by the ISP every time the user logs on, ISPs have logs that link static IP addresses to particular devices and, eventually, to specific users (Bahadur et al., 2002). Due to the need to easily set up servers and remote connections, businesses tend to have static IP addresses while home users are more likely to have a dynamic IP address.[34]

Knowledge of an IP address allows the searcher to obtain other information about a device, service or network. For instance, one can conduct a *trace route*, a computer diagnostic tool for displaying the route (path) of packets across an IP network, in order to find the logical path to the device, which often contains clues to the physical location. Even though identification does not follow automatically from access to an IP address alone, which has led many companies to argue that IP addresses are non-personal data, various other clues can readily be used to identify individuals. Depending on the lookup tools available on the Internet, more information can be revealed including country, region/state, city, latitude/longitude, telephone area code and a location-specific map (Office of the Privacy Commissioner of Canada, 2013). Illustrative of the footprints users leave in cyberspace is the work of Malin et al. (2003), who demonstrated a way to identify a person based on a "trail of seemingly anonymous and homogenous data left across different locations". Using the

[34]In certain cases a static IP address is necessary. For instance,a dynamic IP address should not be used for VOIP, VPN, playing online games or game hosting because it is less reliable than the static IP address and could cause the service to disconnect.

example of an online consumer who leaves the IP address of his computer on every website he visits and combining that with the explicitly identifying information that he provides in order to complete his purchase, the authors explain that "[b]y examining the trails of which IP addresses appeared at which locations in the de-identified data and matching those visit patterns to which customers appeared in the identified customer lists, IP addresses can be related to names and addresses".

4.2.2.3. Potential for re-identification

The digital economy relies on the collection of personal data on an ever-increasing scale as users' information is constantly shared with advertisers, researchers and government agencies. In order to justify the indiscriminate and perpetual sharing and storage of personal data, private companies reassure their customers that their data will only be released in a non-personally identifiable form. The underlying assumption is that personal data is a fixed set of attributes and once data records have been de-identified, either by removing or modifying personal data, they become completely anonymous and thus safe to release, without any possibility of linking them back to individuals. Advances in re-identification techniques expose those promises as too often illusory and reveal the potential risks an individual faces if his data is used in a particular manner (Acquisti & Gross, 2009). A serious flaw in the basic idea behind the distinction between personal data and non- personal data is that it seems to ignore the fact that almost all information, even information rendered anonymous, can become personal when paired with other relevant bits of data. As stated by Ohm (2010), "there is always some piece of information […] that could be combined with anonymised data to reveal private information about an individual". The technical difficulties of de-identifying data raise a challenge to current concepts of personal data. A major problem with defining certain types of information as personal data is that technology increasingly enables the combination of various pieces of non-personal data to produce personal data. As demonstrated by Narayanan and Shmatikov (2009), the term personal data "simply has no technical meaning" because any information that distinguishes one person from another can be used for re-identifying anonymous data; whereas some attributes might be uniquely identifying on their own, any attribute can be identifying in combination with others. The authors show that there is a wide spectrum of human characteristics that enable re-identification and share two key properties: (1) they are reasonably stable across time and

contexts and (2) there is a very small possibility that two individuals are similar due to the large number of the corresponding data attributes.

Another significant factor that renders the permanent de-identification of data unattainable is associated with the exponential increase of personal information both in online and offline records. The phenomenon of data availability, coupled with the development of aggregation techniques, which involve the combination of various pieces of non-personal data to produce personal data, heighten the risks for re-personalisation (Solove, 2008). Ohm (2010) has borrowed the term "accretive" from finance in order to show that the more information about an individual is revealed as a consequence of re-identification, the easier it becomes to identify that individual in the future. When additional pieces of identified information are available, it is highly likely that there will be more elements in common between the de-identified and already identified data, which leads to the linking of the former with the latter. That is why, even though re-identification algorithms are more difficult to implement and more computationally expensive compared to re-identification methods based on demographic records, the former kind of re-identification is a one-time effort and it can be applied on "thousands or even millions of individuals" (Narayanan & Shmatikov, 2010). Current corporate practices permit this linking as they play a significant role in shaping the amount and types of information that are available online by deploying systems which track users' activities often without their knowledge (e.g. Google Buzz, Facebook Beacon)[35] (Federal Trade Commission, 2010).

4.2.2.4. Re-identification cases

The following re-identification cases, where sophisticated data administrators placed unjustified faith in anonymisation, exemplify the vulnerability of supposedly de-identified datasets to re-identification and highlight the perils of anonymisation in light of recent advances in re-identification techniques. As indicated below, a single piece of non-personal data does not exist alone; rather, data form part of a landscape where extensive information is

[35] *See*, for example, McCarthy, C., 2007, Facebook's Zuckerburg: "We Simply Did a Bad Job" Handling Beacon, *CNET*, December 5, available at http://news.cnet.com/8301-13577_
3-9829526-36.html; FTC, 2011, FTC Charges Deceptive Privacy Practices in Googles Rollout of Its Buzz Social Network, March 30, available at https://www.ftc.gov/news-events/press-releases/2011/03/ftc-charges-deceptive-privacy-practices-googles-rollout-its-buzz.

available and when such information is aggregated, it can produce even more information in a way that renders de-identification particularly difficult.

4.2.2.4.1. Sweeney's study

As early as in 2000, a study conducted in the U.S. by Latanya Sweeney demonstrated that the combination of information relating to an individual, such as the gender, the date of birth and a ZIP code could lead to the identification of the vast majority of individuals in the U.S. (Sweeney, 2000). By processing 1990 pieces of data, which were not regarded intimate or particularly sensitive, Sweeney discovered that 87% of individuals in the U.S. were uniquely identified. According to the study, individuals could be identified based on even less specific information; for instance, information such as gender, date of birth and the city could reveal the identity of 53% of American citizens, while 18% could be identified by the combination of their gender, date of birth and country. Sweeney's study followed her successful attempt in the mid-1990's to find the health records of the then Governor of a government agency called Group Insurance Commission, which decided to disclose the records summarising every state employee's hospital visits to any researcher that would request them, based on the belief that data remains protected insofar as fields containing name, address, social security number, and "other explicit identifiers" have been removed.

4.2.2.4.2. AOL search leak[36]

In 2006, AOL released to the public a file containing twenty million search queries used by 650,000 users of the company's search engine summarising three months of activity in order to embrace the vision of an open research community. Although search queries are the kind of information that are usually treated as a closely guarded secret, AOL had tried to anonymise the search-query data before releasing; in the abstract, anonymised queries appear to be non-personal data. To this end, the company suppressed any obviously identifying information, such as the username and the IP address, and replaced them with unique identification numbers so that the data remains useful for research purposes. However, AOL had mistakenly believed that the information released was truly anonymous since, as it was

[36] *See*, e.g., Arrington, M., 2006, AOL Proudly Releases Massive Amounts of Private Data, August 6, available at http://www.techcrunch.com/2006/08/06/aol-proudly-releasesmassive-amounts-of-user-search-data; Zeller, T., 2006, AOL Executive Quits After Posting of Search Data, *N.Y. TIMES*, August 22, available at http://www.nytimes.com/2006/08/22/technology/22iht-aol.2558731.html.

proved in the days following the release, personally identifiable information was available in many of the searches making it possible to identify an individual and their history.

4.2.2.4.3. Narayanan-Shmatikov's studies

Two months after the AOL's release of search queries, Netflix made a database of movie ratings publicly available as part of a contest to improve the predictive capabilities of its movie-recommending software. Before disclosing one hundred million records, which contained the movie rated, the rating assigned and the date of the rating, Netflix tried to anonymise them by removing identifying information such as usernames and assigning a unique user identifier to preserve rating-to-rating continuity. Two weeks after the release, two computer scientists from the University of Texas, Arvind Narayanan and Vitaly Shmatikov found a way to link the data released with the movie ratings that participating individuals gave to movies in IMDb, a movie-related website that posted such ratings publicly on its website, and showed how easily an individual can be re-identified in a database. For instance, their research concluded that if an adversary knew approximately when an individual in the database had rated six movies, he could identify the individual 99% of the time (Narayanan & Shmatikov, 2008). Another re-identification study conducted by Narayanan and Shmatikov (2009) demonstrated the feasibility of large-scale re-identification using the local structure of social networks. In this study, the researchers anonymised the entire Twitter graph, which resulted in nameless, identity-free nodes representing users connected to other nodes representing Twitter 'follow' relationships, and then they compared that stripped-down graph to public data harvested from the Flickr website. By comparing the structure of the anonymised Twitter's 'follow' graph and the non-anonymised Flickr's 'contact' graph, the researchers managed to identify the usernames or full names of one third of the people who subscribed to both social networks, without using any single piece of identifying information such as usernames, activity information or photos.

4.2.2.5. Personal data as a constantly changing concept

Computer science proves that initially non-personal data can be transformed into personal data when linked to individuals, thereby the line between personal and non-personal data is not fixed but rather depends upon the constantly changing technologies. Ohm (2010) questions whether it is even possible to maintain such a distinction positing that personal data is an "ever-expanding category" that "will never stop growing until it includes everything".

In fact, given the increase in computational power and the versatility of re-identification algorithms, the current distinction is difficult to maintain since today's non-personal data might be tomorrow's personal data. Illustrative of the changing landscape of technology is the report of the Privacy Protection Study Commission conducted as early as 1977, according to which "[a] major problem created by the widespread adoption of computer and telecommunications technology to personal-data record keeping is the inability to anticipate and control future use of information" (Privacy Protection Study Commission,1977). The report pointed out that record systems were designed and modified in a way that merely covered the immediate and specific needs, without taking into consideration the long-term implications of the computerisation of other areas of record-keeping. The fact that the re-identification of personal data depends on the current state of technology renders identifiability a complex concept that cannot be determined *a priori* but is rather driven by context.

The emergence of advanced re-identification technologies, the availability of the personal data of millions of individuals and the increasing economic incentives for potential attackers illustrate not just the limitations of anonymisation techniques, but also "the fundamental inadequacy" of the current data protection paradigm, where the distinction between personal data and non-personal data is always made in abstract. Instead, given that the risk of re-identification changes over time, the data protection regime should be built on a case-by-case basis (Narayanan & Shmatikov, 2010). To this end, organisations processing personal data should not assume that data that is anonymous now will remain anonymous in the future, but they should regularly perform a thorough assessment of the re-identification risks and carry out a periodic review of their policies on the release of data and the techniques used to anonymise it, taking account of "all the relevant contextual elements", such as the appeal of the data for targeted attacks (the nature of the data and the sensitivity of information), the availability of public information resources, the envisaged release of data to third parties etc. (Article 29 WP, 2014).

4.2.2.6. Should the concept of personal data be abandoned or not?

The amenable nature of what constitutes personal data, the fragility of anonymisation techniques and the power of re-identification technologies have led some commentators to suggest that the concept of personal data as the foundation of the EU data protection

framework should be entirely abandoned. As demonstrated above, simply removing personal data should no longer be considered sufficient enough to provide meaningful guarantees of data protection and therefore the distinction relied on whether particular data types are more linkable to identity than others seems to have lost its scientific basis. It has been argued that the centre of focus of the data protection regime should be now placed to new ways of evaluating the risks to data protection, able to embrace the future contexts. It is noteworthy that despite the aforementioned re-identification incidents and studies, which proved that information (such as movie ratings or search queries) that were not categorised as personal data in the past can now reveal one's identity, there has not been any change in the law addressing the processing of such data, while companies continue disclosing this kind of information connected to sensitive data in supposedly anonymised databases.

A different approach has been suggested by Schwartz and Solove (2011), who posit that the concept of personal data can still be functional inasmuch as it is modified according to "a risk-based continuum". The new standard-based model proposed contains the following three categories of information, that is, information that refers to: (a) an identified individual, an individual whose identity is "ascertained" or can be "distinguished" from the group; (b) an identifiable person, when the re-identification risk is low or moderate; (c) a non-identifiable individual, when the re-identification risk is remote. To classify the information to the appropriate category, account should be taken of "the means likely to be used by parties with current or probable access to the information, as well as the additional data upon which they can draw" and other contextual elements, such as "the lifetime for which information is to be stored, the likelihood of future development of relevant technology, and parties' incentives to link identifiable data to a specific person".

Another approach (Essayas, 2015) that looks at the re-identification risk threshold, in order to examine whether data should be considered anonymised, and thus exempt from the data protection provisions, suggests a model where different layers of risks of harm should be established and consideration should be given to the sensitivity of the data and the context of its usage. The layers include the following four cases: (a) anonymisation is applied to sensitive data and the data would be publicly available; (b) anonymisation is applied to sensitive data and the data would be available with limited access; (c) anonymisation is applied to non-sensitive data and the data would be publicly available; (d) anonymisation is applied to non-sensitive data and the data would be available with limited access. The author

suggests that – at least - the fourth layer be subject to less strict requirements for anonymisation.

Regardless of the approach suggested with respect to the concept of personal data, most commentators agree on the idea of abolishing the term ‘anonymisation’ as misleading because it overpromises and thus creates a false sense of security that does not correspond to the actual outcome. The inherent residual re-identification risk linked to any technical and organisational security measure dictates the need for a more technically accurate term which connotes the effort to achieve anonymity and not presenting the achievement as accomplished; terms such as ‘anonymisation technique’, ‘pseudonymisation’, ‘de-identification’ and ‘scrub’ have been suggested instead.

4.3. Technological Assumptions Hidden in the Definition of Personal Data Processing

The definition of the processing of personal data, as described in Article 2(b) of the DPD, begins with a broad statement according to which the term ‘processing’ includes “any operation or set of operations which is performed upon personal data, whether or not by automatic means”, and is accompanied by a list of actions with respect to personal data that could be considered as processing activities, “such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction” of personal data. By being function-based, this broad opening definition can easily accommodate future technologies and hence it seems to be in accordance with the requirement for technological neutrality since the processing activities that any of those technologies would entail will either fall within one of the examples listed in the DPD or within the general ‘use’ of personal data. In addition, the definition cannot be criticised for limiting the possible processing activities to those explicitly mentioned as the above list should be regarded as merely illustrative and not restricting. As a matter of fact, the technological neutrality of this definition has already been tested with the development of – different in nature - technologies that emerged after the adoption of the DPD and prove the sustainability of the DPD. The way in which geo-location activities, for example, function

enables the collection, recording, organisation and storage of data which can then be retrieved, used, disclosed, made available, blocked or erased.[37] Another example can be taken from the radio-frequency identification (RFID) systems that use tags in order to collect, record and store data to enable the data's retrieval, use and disclosure.[38] And the list of examples of current technologies that process personal data performing operations that fall within the above definition is inexhaustible.

Even though the legal scope of the definition is written in a technologically neutral manner, carefully reading the list of the possible operations performed upon personal data can reveal the underlying technological assumptions of the definition. The order in which those activities are mentioned is indicative of the linear sequence in which the above list has been organised as the activities seem to follow a certain chronological sequence. To make this temporal linearity assumption clearly understood, Birnhack (2013) has classified the processing activities into the following clusters: collection and recording describe the *input* of the data; organisation, storage and adaptation refer to the *management* of the database; retrieval, consultation and use refer to the *internal usage*; disclosure, dissemination, making available, alignment, combination (and possibly blocking) refer to *the output*; erasure and destruction describe what happens with the data once it is no longer in use and thus they refer to the *clean-up* of the data. The progressive assumption reflected in the manner the list of activities has been organised contains the following two assumptions. First, the steps that the processing of the data follows can be likened to the way the life of a human being has been conceived; it is first born, then it grows up, it becomes productive, and it ultimately dies. In this analogy, database appears to play the role of a human being's 'home'. Similarly, the linear sequence assumes a particular technological environment, in which data is first collected and, once stored in a database, it can be used or transferred to third parties. Second, the central processing activities seem to be the internal and external usage (input and output) of data, whose function is served by the rest activities in such a way that the first step of input is seen as a preparatory stage and the final steps (clean up) as a wrapping up mechanism.

Moreover, the assumption about linearity indicates that there are certain players involved in each segment of this linear structure. In particular, the initial segment of the input of the data

[37]*See* Art. 29 DPWP Opinion 13/2011 on Geolocation Services on Smart Mobile Devices, WP 185, May 16, available at http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2011/wp185_en.pdf.
[38]*See* Art. 29 DPWP, 2005, Working Document on Data Protection Issues Related to RFID Technology, WP 105 January 19, available at http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2005/wp 105_en.pdf.

involves the data subject and the data controller (and possibly the data processor acting on behalf of the former); the main player in the segments of management and internal usage is the data controller, perhaps with the assistance of the data processor; the output segment is under the direction of the data controller and the recipient of the data; the clean-up of the data is exclusively driven by the data controller (Birnhack, 2013). It is noteworthy that, according to the above structure, the data subject appears to have a role in the processing of her personal data merely in the first stage when she is asked to consent to or decline the processing and a minor role in the last three stages where the DPD provides the data subject with the rights of access and rectification thus enabling her some power to assure that her rights are not infringed. To compensate for the limited control that the data subject can exercise over her data, the DPD has introduced mechanisms for extending that control beyond the first stage by imposing further duties to data controllers, such as keeping the data in a form which permits identification of data subjects for no longer than is necessary (Article 6(1)(e)), processing data fairly, lawfully and for legitimate purposes (Article 6(1)(a)(b)), maintaining confidentiality and security (Articles 16, 17). Nevertheless, the fact that the data subject can take active part solely in the first stage, coupled with the likelihood that in practice data controllers may not always fulfil their duties in a way that respects the data subject's rights, demonstrate the problematic nature of this structure with respect to the data subject's control over the processing of her personal data. In this respect, two new elements introduced by the GDPR, the proposed principle of accountability as well as the right to be forgotten, which would enable individuals to demand the erasure of their data under certain circumstances, can be seen as a means of strengthening the power of data subjects (Recital 33 and Article 8).[39]

Even that limited power granted to the data subject in the form of informed, explicit and unambiguous consent, which the DPD has much relied upon, appears to be challenged by the emerging technological trends. The DPD takes as granted that data subjects are given the opportunity to make an informed decision with regard to the collection and processing of their personal data. However, new technologies collecting data in a way that bypasses the initial segment of the linear structure, namely the meeting point between the data subject and the data controller, deprive the former of exercising meaningful control. More and more data is nowadays collected in various settings and in less detectable ways (Commission of the European Communities, 2003). When technologies such as cloud computing, which is

[39]The GDPR suggests that in the case of "significant imbalance between the position of the data subject and the controller," consent cannot provide a legal basis for processing (Article 7(4)).

increasingly gaining ground, are deployed, data is no longer stored in the user's hard disk, which, first, indicates the limitations of the DPD's linearity assumption and, second, results in data subjects exercising less control over their data (Article 29 WP & Working Party on Police and Justice, 2009; Gellman, 2009). Frequent are also the cases where the data subject can hardly control some categories of data about herself, such as biometric, genetic, cognitive data or data related to her behaviour (PRACTIS, 2011). Industry's practices based on tacit data, such as the example of targeted online behavioural advertisement, are gaining market power while revealing the absence of the users' physical ability to exercise control over their personal data (Tene & Polotensky, 2012).

As explained above, the DPD assumes a technological environment where the powerful player is the data controller; in other words, the environment of Web 1.0 and the related Privacy 1.0.[40] In the Web 2.0 environment, however, users of social networks publish and process personal data of other users – or even of other users' 'friends' - in such a way that the role of the operator of a social network, in its current form, is restricted to technical control without him having control over the content of a user's data processed by other users (Edwards & Brown, 2009; Grimmelmann, 2009). The technological and social assumptions of the DPD fail to address the way in which social networking systems operate as they merely address the relationship between the end-user and the operator of the social network. As a matter of fact, the Article 29 WP (Opinion 5/2009) concluded that the providers of social networking services should be considered data controllers and that most users act for purely personal or household purposes which are exempt from the DPD. Even though the Opinion clarifies the rights of the data subjects and the obligations of the data controllers, in terms of the users' control over their data during their interaction with other users, the Opinion adopts an advisory approach encouraging users to first obtain the consent of their peers before uploading information related to them, without mentioning any legal safeguards. Given the increase of this kind of interaction in social networking platforms, the lack of security and confidentiality safeguards questions whether the DPD is appropriately equipped to deal with this new source of threat to individuals' privacy.

[40] *See* Zittrain, J.L., 2008, *The Future of the Internet and How to Stop It*, London: Penguin Group.

4.4. Conclusion

Notwithstanding the fact that the definitions included in the DPD are technology neutral in terms of the language and references used, the above analysis has demonstrated the technological assumptions upon which the DPD hinges. Most importantly, the faith in robust anonymisation, which has thoroughly infiltrated the EU data protection system, involves the danger of treating data likely to be re-identified as anonymous and therefore exempt it from the ambit of the data protection regime. The emergence of re-identification techniques able to manipulate databases and link the data to the identity of an individual renders the irreversible anonymisation assumption obsolete and reveals the inadequacy of the current data protection paradigm. The 'all-or-nothing' approach, according to which data is either anonymised or not, appears to contradict with the concept of negligibility prevailing in engineering that suggests that in practice there is no zero-risk situation and thereby data should be considered anonymised in case of a low identification risk. The danger involved in such an assumption is that the definition should be understood within a certain technological paradigm and hence it is sufficient as long as the paradigm remains the same. Once, however, the paradigm changes, the concept of non-identification will collapse and lawmakers will need to find new ways to regain the lost balance. Currently, legal perspectives on anonymisation are often solely concerned with the question of whether a certain technique meets the criteria for providing a safe harbour from the entire application of data protection rules without taking into consideration the risks involved in the implementation of such measures. In spite of the increasing number of re-identification incidents, there has not been any amendment in law addressing the processing of such data, while companies keep collecting, using and sharing (often sensitive) personal data included in supposedly anonymised databases. Furthermore, the DPD assumes a technological environment where the data controller is the powerful player, thus failing to regulate social networking environments and the rapid expansion of Big data. Given the dynamic nature of re-identification risks, which increase over time, it is essential that the law recognises the fact that data identifiability cannot be determined a priori but is rather driven by context, and therefore the line between personal and non-personal data is not fixed but depends on the constantly changing technologies.

CHAPTER 5: Hidden Technological Assumptions and Hazards of Technological Neutrality

5.1. Introduction

The widely accepted view shared among legislators that the enactment of technology neutral laws is the most suitable means of guaranteeing the sustainability of the laws while granting flexibility to the private sector often leads them to either ignore or fail to foresee the implications involved in certain technology neutral laws. Another issue arising from technology neutral laws is associated with the failure of legislators to achieve their initial goal of drafting laws that do not favour any particular technological design. The chapter begins with a brief description of the prevalent legal approaches to regulating electronic signatures whose divergences stem from the debate over the most effective manner to better address the use of electronic signatures. The different approaches focus on the question of which need should be prioritised in the context of electronic signatures - the need for promoting innovation or providing security safeguards. Although the two-tiered approach adopted by the EU legislator appears, at first glance, to combine both the technology neutral and technology specific approaches, thus combining the benefits of both, examining more thoroughly the EU Directive regulating electronic signatures reveals that technology favourism undermines its purportedly technology neutral nature. In the first place, the law seems to provide for legal acceptability of electronic signatures on a non-discriminatory basis aspiring to embrace all future technologies. Yet, the requirements laid down for the category of advanced electronic signatures demonstrate that certain technologies are favoured by being afforded special presumptions, in contrast to the legislator's initial aim to avoid stipulating requirements that could only be met by specific technologies. The chapter proceeds with the concerns stemming from technology neutral laws in the field of surveillance and data processing activities, where technical considerations with severe practical implications are not included in primary legislation but are rather left to secondary legislation, thus reducing democratic protections. Drawing on the example of laws whose technological ambiguity involves the danger of excessive powers being granted without due process, it becomes clear that the technology neutral approach should not disregard considerations relating to activities that jeopardise individuals' fundamental rights when establishing the appropriate legal safeguards.

5.2. Technological Neutrality and Electronic Signatures

5.2.1. Three categories of legislative approaches to electronic authentication

Electronic signatures often appear in discussions regarding technology neutral policies. The need for any legal approach to electronic authentication to balance the inherent tension between technological neutrality and technological specificity has fuelled much scholarly debate since the emergence of electronic signature laws. The debate revolves around the question of whether the laws regulating electronic signatures should remain neutral towards technology or prescribe specific legal consequences for the use of electronic authentication systems. The central source of disagreement is related to the potential for innovation that is promoted by technological neutrality, on the one hand, and the security safeguards which a technology specific legislation would provide. While the proponents of technological neutrality argue that it *"fosters innovation, whereas regulating specific technologies inhibits the marketplace from continuing to develop more effective technologies than current e-sign technology"* (Wright & Winn, 1999), the counterargument used by the opponents of such legislation is that technology specific laws can *"boost the security industry forward, creating and standardizing the most secure and viable technology that exists"* (Koger, 2001).The ways in which lawmakers have sought to accommodate the aforementioned conflicting concerns can be grouped into the following three categories.

5.2.1.1. The prescriptive approach

Following the example of the first electronic signature legislation of the State of Utah, the earliest countries to adopt such legislation endorsed the prescriptive approach mandating specific technology – digital signatures – that is, signatures based on public-key cryptography (Baker & Yeo, 1999). Although electronic authentication can be achieved in many ways, for instance, with PIN codes, pass phrases or biometrics, at the time when the first electronic signature initiatives began to emerge, the use of asymmetric or public key as a means of creating digital signatures was the most widely used and universally accepted means for securely authenticating electronic documents; therefore, early initiatives were limited to promoting and facilitating digital signature technology (Barofsky, 2000).

5.2.1.2. The minimalist or functionalist approach

Most common laws jurisdictions of the world initially adopted the technology neutral and market-oriented approach, according to which no detailed technical standards regarding the use of different authentication techniques should be imposed and thus no special presumptions were limited to PKI (public key infrastructure) or other particular technologies (Fischer, 2001). Legal recognition is granted to any electronic signature scheme as long as it meets certain general criteria that prove its capability to providing authentication (Koops, 2006). For example, the U.S. E-Sign Act 2000 provided that no electronic signatures of whatever type may be "denied legal effect, validity, or enforceability solely because it is in electronic form".[41]

5.2.1.3. The two-tier or hybrid approach

The two-tier approach is founded on a policy of limited technological neutrality but constitutes a more market-driven legislative model compared to the prescriptive one (Fischer, 2001). More specifically, at the first level, this approach grants a basic set of legal benefits to all electronic signatures, while at the second level, a broader array of benefits is provided to a class of approved technologies usually referred to as 'secure' or 'qualified' technologies, which should meet either general criteria or criteria relating to the specific techniques of asymmetric cryptography. As a result, documents authenticated by one or more of the above technologies are entitled to a more robust set of legal presumptions, such as a presumption concerning the identity of the signer or the integrity of the document's content (Baker & Yeo, 1999).

5.2.2. Should laws governing electronic signatures be technology neutral?

Laws regulating digital signatures have sparked much criticism (Biddle, 1997; Robertson, 1998; Boss, 1999; Wright & Winn, 1999). The opponents of the prescriptive approach contend that by giving legal recognition only to one type of technology, there is a risk that technological improvements may be stymied (Electronic Commerce Expert Group, 1998). Without regard to market-oriented solutions, "favouritism" toward digital signatures is likely to prevent other superior technologies from entering the marketplace (Wright & Winn, 1999;

[41] 15 U.S. Code § 7001 - General rule of validity

International Working Group On Electronic Authentication, 1999). Given the pace of technological development, the law must change every time new technologies arise that may be considered more advanced. Therefore, the role of the law should be limited to dealing with the legal effects of electronic signatures, while its goal is to *"remove the barriers to electronic commerce, treat electronic communications on a par with paper communications, and not to favour one technology over another (technology neutrality) nor one business model over another (implementation neutrality)"* (Boss, 1999). Based on the argument that different technologies may be preferable for different purposes, the proponents of the minimalist approach argue that the market should be left free to choose the technology and implementation scheme that best suit their particular transactions and also determine issues such as the level of security and reliability required for electronic authentication.

Contrary to the prescriptive approach, which seems to give undue benefits to a technology that is in the earliest stages of commercial use, as was the case of digital signatures when the first electronic signature laws were enacted, the primary advantage of the minimalist approach lies in the fact that it leaves room for the development of other authentication mechanisms and thus satisfies the general desire of legislators to avoid the rapid obsolescence of new laws (Biddle, 1997; Corwin, 1998). What is more, countries adopting the minimalist approach are less likely to be left outside the mainstream of technological developments internationally, which is regarded of vital significance given the lack of any internationally uniform legislative approach. For all the reasons described above, technological neutrality has become an increasingly prevalent option as a legislative approach to regulating electronic authentication (Baker & Yeo, 1999).

Despite the fact that technology neutral laws are viewed as more appropriate in the field of electronic authentication, they are not without limitations. The minimalist approach is often criticised for being too vague and thus creating too much legal uncertainty (Fischer, 2001). By granting legal recognition to all authentication techniques, technology neutral laws provide little guidance as to the criteria that an authentication means should meet in order to be considered adequately secure and hence such laws do not provide a satisfactory answer to the question of legal recognition (Koger, 2001; Koops, 2006). Given that legal certainty is key to stimulating widespread public trust in electronic transactions, a minimalist approach can have harmful effects on e-commerce. The absence of legally recognised security procedures may result in potential participants abstaining from engaging in e-commerce,

while the possible use of insecure techniques may give rise to fraudulent activities, further decreasing confidence in e-commerce (Robertson, 1998; Boss, 1999).

Based on the argument that not all technologies necessarily require the same type of legally defined trust infrastructure or may be afforded the same presumption of security and integrity, the opponents of the minimalist approach claim that it fails to establish a reliable security infrastructure (Baker & Yeo, 1999). A known and reliable authentication mechanism with established legal consequences is regarded a necessary prerequisite for the continued expansion of e-commerce, which can merely be achieved through detailed technical standards (Kuner, 1999). However, it is difficult for legislators to accord specific and meaningful legal consequences to the use of electronic authentication techniques that are not yet even conceived. On the contrary, the security and reliability of already known techniques enable lawmakers to grant greater legal benefits to the use of those techniques. That was the reason why, when the first electronic authentication laws appeared, the supporters of the "technology movement" argued that legal recognition should only be limited to digital signatures (Kuner, 1998; Blythe, 2005)). At the time, digital signatures were seen as the only technical concept that could satisfy the high technical security standards required in e-commerce by guaranteeing authenticity and integrity. As stated by the CEO of a U.S. Certification Authority, digital signatures are "*security tools that are backed by stringent policies and procedures allowing people to trust the authenticity and enforceability of electronic transactions".*[42]

5.2.3. Technological neutrality of the EU legal framework surrounding the electronic signatures

After several European countries began to independently enact electronic signature laws in the late 1990s, the EU issued the Directive on a Community Framework for Electronic Signatures (e-Signatures Directive)[43] in 1999. In order to reach a middle ground between technology neutral and technology specific approaches, the Directive adopted a two-tiered approach in defining the legal effects of an electronic signature. On the one hand, it provides

[42] Press Release, Digital Signature Trust, Digital Signature Trust Becomes Licensed as Certification Authority in Texas (May 16, 2001) at <http://www.prnewswire.com/news-releases/digital-signature-trust-becomes-a-licensed-certificate-authority-in-three-new-states-71831002.html>

[43] Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures

legal acceptability of electronic signatures in general on a non-discriminatory basis.[44] On the other hand, the Directive distinguishes between basic and advanced electronic signatures, with the latter being admissible in legal proceedings and requiring a greater level of security than the former. An 'advanced' electronic signature should be based upon a 'qualified certificate' – a certificate that meets specific security standards – issued by a qualified 'certification-service-provider' and generated using a 'secure-signature-creation device'.[45] The specific requirements for a qualified certificate, certification-service-provider and 'secure-signature-creation are laid down in the Annexes of the Directive.

As indicated in Recital 8, according to which *"rapid technological development and the global character of the Internet necessitate an approach which is open to various technologies and services capable of authenticating data electronically"*, the legislators' aim was to maintain both the elements of a technology neutral approach – sustainability and innovation. The advantage of the hybrid approach adopted by the EU Directive is that it combines the benefits of both the aforementioned approaches. Not only does it ensure a level of legal certainty, which is necessary to build and maintain trust in e-commerce, but it also appears to be flexible and adaptable to new technological developments (Fischer, 2001). However, critics of this approach argue that it involves excessive government regulation as it seems to overprotect certain technologies at the expense of innovation and hence it does not permit sufficient breathing room for market forces (Baker & Yeo, 1999). As a matter of fact, only the first level of the Directive accepts electronic signatures on a technology neutral basis. The approach taken by the Directive has been criticised for promoting 'technological favouritism' on the grounds that if the criteria listed in the Annexes were met, then certain favoured technologies would be afforded special presumptions, such as the presumption of authenticity (Barofsky, 2000). At the time of the Directive's drafting, that presumption could be limited functionally to digital signatures as qualified certificates were unique to PKI technology. Even though the Directive is at first level concerned with the function of authentication technologies, not with specific technologies, its technology specific approach is illustrated in the requirements for advanced electronic signatures, which are closely related to the technology of digital signatures, and thus the Directive's emphasis on attainment of security does seem to implicitly support the use of more sophisticated and security-minded technologies such as PKI (Wright & Winn, 1999; Blythe, 2005; Koops, 2006). Therefore,

[44] *See* Article 5
[45] *See* Articles 2, 5

otherwise ‘secure’ authentication mechanisms that are not based on a ‘qualified certificate’ or created by a ‘secure-signature-creation device’ appear to fall short of presumptive validity and would not enjoy equality with digital signatures (Corwin, 1998). By giving technology-specific rules for a certain type of electronic signatures, the Directive seems to have missed its target to be technology neutral.

Although the legislators of the EU Signature Directive intended to draft a technology neutral law able to embrace all the future technologies, fifteen years later the Directive was replaced by a new legislation – the eIDAS Regulation[46] – on the grounds that, amongst others, the Directive did not cover new technologies having emerged since 1999, such as mobile or cloud signing (European Commission, 2014). The eIDAS Regulation follows the same approach to be technology neutral by claiming to avoid requirements that could only be met by a specific technology, thus avoiding precluding any of the existing or emerging technologies (ABC4Trust, 2013; European Commission, 2014). To this end, the Regulation is broader in scope as it does not only regulate electronic signatures but contains EU-wide rules concerning trust services, such as the creation and verification of electronic seals, electronic time stamps, registered electronic delivery services, as well as the creation and validation for website authentication. However, little has changed as regards the electronic signatures under eIDAS compared to the e-Signature Directive. Apart from the terminology that has been amended – a ‘secure signature creation device’ becomes a ‘qualified signature creation device’ – and the concept of ‘qualified electronic signature’ that has been introduced[47], the definitions largely match. It is noteworthy that the security requirements for electronic signatures are essentially similar to those of the Directive though clarified and laid down in more detail[48] (Livesey, 2012; Brazell, 2016). The same technology specific approach seems to have been adopted with respect to all trust services as more technical standards relating to every aspect of the certification and qualification processes, including preservation (archiving) services, have been introduced.[49]

[46] Regulation No910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC
[47] See Article 3
[48] See Articles 26, 29, 30, 32, 33, 34
[49] See “eIdAS – European Council adopts electronic identification and trust services regulation”, 14 July 2014, at <http://certifiedsignature.eu/2014/07/24/eidas-council-adopts-electronic-identification-rules/>.

5.3. Technological Neutrality and Surveillance Laws

5.3.1. Should laws governing surveillance activities be technology neutral?

As is the case with all technology-related laws, when drafting surveillance laws legislators encounter the dilemma of choosing between technology neutral and technology specific language. In the majority of the cases, the great challenge of lawmaking to keep up with the latest advances in technology results in legislators embracing technological neutrality as a principle, which should only be infringed for exceptional reasons (Sales, 2009). By writing technology neutral provisions that only describe the general characteristics, the effects and the purposes of a technology without referring to the specific features or functions of a particular type of technology, legislators manage to keep the law open to new technological developments. Those who espouse technological neutrality argue that technology specific provisions become not only anachronistic but also under-inclusive as time goes by, since the targeted technology will either evolve into a new form or be replaced by new technologies, none of which will be covered by technology specific provisions (Reed, 2007). Furthermore, technological neutrality also seems to satisfy the need for consistency, that is, the need to treat similar technologies alike without leaving certain technologies unregulated (Van der Haar, 2007). Another argument in favour of technological neutrality is associated with the recognition of institutional shortcomings. Legislators who draft high-level laws are not necessarily equipped to understand complicated technological details, which are thoroughly analysed by further guidelines and implementing acts (Berkowitz, 2007).

Although technological neutrality may be the most appropriate choice in certain contexts, the blind adherence to this approach can generate undesirable effects or even involve hidden hazards in the surveillance context. A close examination of the arguments supporting technological neutrality reveals some underappreciated logical flaws and gaps (Moses, 2007; Ohm, 2010). For instance, by trying to avoid the aforementioned problem of under-inclusiveness technology neutral provisions often tend to reach the other extreme permitting over-inclusiveness due to the open-textured and vague terms used in such provisions, which leave wide margins for interpretation. Besides, when a law expands over time, the grounds on which it was initially based cannot always remain tenable under different circumstances (Ohm, 2010). This holds true, for example, with regard to the way a technology impacts on the right to privacy; over the course of the years, changing technology brings new challenges

to privacy.[50] By applying the same rules to any technology developed in the future, irrespective of how it operates or where it is deployed, technology neutral laws may refrain from discriminating against certain technologies, but in this way they fail to take into account the specific characteristics or effects of certain technologies that need to be treated differently (Kerr, 2004; Solove, 2004; Nissenbaum, 2009). The proponents of technological specificity argue that legislators should first study a specific technology and then tailor the law to the idiosyncrasies of the specific context in order to avoid future unforeseen effects, especially when specific new forms of technology enable both new methods for committing crime and new forms of surveillance (Kerr, 2004; Ohm, 2010). As illustrated in the work of Balkin & Levinson (2006), "new technologies of surveillance, data storage, and computation" have given rise to the so-called "National Surveillance State", where the amount of intelligence and surveillance the government conducts in the name of protecting national security has significantly increased.

Contrary to technology neutral surveillance laws that deprive the legislature of the power to exert oversight over the surveillance activities of the executive, technology specific laws force the executive to consult with legislators in the case of significant technological amendments, leading to a more participatory democratic oversight to the conduct of national surveillance, able to prevent possible surveillance abuses or excesses (Ohm, 2010). Another underappreciated benefit that technology specific laws serve, and which is commonly presented as a downside, is the fact that their validity lasts the same amount of time as the technologies they regulate. For example, a law that governs only the use of the telephone will not also govern the use of the Internet. Once new technologies are introduced, a technology specific law needs to be replaced. In this regard, technology specific laws seem to operate as traditional sunset provisions, that is, provisions within a law that cease to have effects after a specific date, unless legislative action is taken to extend them. What is more, technology specific laws seem also to satisfy the reasons for which legislators enact sunset provisions. According to Jacob Gersen (2007), the aim of sunset provisions is to offset asymmetries, reduce error costs in the face of uncertainty and correct limits of cognitive bias; likewise, the role of technology specific laws lies in clarifying the doubt and uncertainty regarding the evolution of a specific technology. More importantly, technology specific laws do not expire

[50]See below the privacy concerns related to technological neutrality when applied to traffic data.

after a fixed period but whenever technological changes call for a re-evaluation of the circumstances that first led to their enactment.

5.3.2. Hazards of adopting a technology neutral approach to regulating lawful access to traffic data

The traditional technology-related laws were established with traditional technological environments in mind, such as the plain old telephone systems (POTS). This is reflected, for example, in the words of Lord Taylor of Holbeach, who commented on the Regulation of Investigatory Powers Act (RIPA) 2000, as follows: *"Much of it relates to fixed line or mobile telephony, so it also covers web-based email and social media communications"* (House of Lords, 2014). Similar approach was adopted by the Canadian government while ratifying the Convention on Cybercrime 2001, when it suggested that all telecommunications services should be treated alike based on the fact that *"the standard for Internet traffic data should be more in line with that required for telephone records and dial number recorders in light of the lower expectation of privacy in a telephone number or Internet address, as opposed to the content of a communication"* (Government of Canada, 2002). In the days of POTS, there was a clear distinction between the content of communications and traffic data (e.g. calling numbers, numbers called, call duration etc.), which was considered derivative and did not require strong legal protection. On the contrary, the content of communications was regarded sensitive and therefore any breach of confidentiality required constraints, such as politician-authorised warrants in the UK or judicial warrants in the U.S. Since 2000, governments have been updating their legislative frameworks to apply to modern communications infrastructures, but there has been widely noticed a trend towards the use of ambiguous terminology that barely takes into account the challenges that new technologies bring about. Instead of acknowledging that changing technological environments "alter the habitat of a policy" and hence new policies need to reflect the specific characteristics of the new environment, the policy language developed under the POTS has been sustained through technology neutral laws (Escudero-Pascual & Hosein, 2004).

The dangerous consequences of applying old protections to new communications infrastructures are clearly illustrated in the case of the investigative powers of access to traffic data and, more specifically, in the way the latter has been treated by the law. As demonstrated in the work of Escudero-Pascual & Hosein, who used a set of data records from different

communications infrastructures in order to show the varying levels of details that can be derived from that data that can be accessed by law enforcement powers, "traffic data's constitution differs by communications medium" (Escudero-Pascual & Hosein, 2002). Given that the more sophisticated a communications medium is, the more information can be deduced by traffic data, updating legal definitions of traffic data ignoring the technical details and the increased sensitivity of such data, in the name of technological neutrality, seems to be quite problematic. If different data can be obtained when shifting between infrastructures, one can imagine the extent to which the problem is exacerbated in the case of converging infrastructures (e.g. mobile communications systems magnify the sensitivity of traffic data). Analysing the privacy concerns of technological neutrality when applied to traffic data, Escudero-Pascual & Hosein argue that this approach is likely not only to provide access to a greater amount of data under a weak regime of protection, but may also "create a shift in the perception of the process of collection"; the data collected depends on both the infrastructure and the means of collection. The quest of technological neutrality involves the risk of surveillance laws also becoming 'content-neutral' focusing on whether the information gathered can be obtained by a certain technology instead of examining whether that information is content or not.

The Council of Europe has acknowledged that governments should take into consideration the high sensitivity of traffic data when establishing the appropriate legal safeguards because of the breadth of traffic data whose collection may "permit the compilation of a profile of a person's interests, associates and social context" (Council of Europe, 2001). In fact, traffic data analysis generates more and more sensitive information about an individual's interactions, often disclosing as much as would be discernible from the content of communications. The increased volume of communications, coupled with the increased ability of the law enforcement agencies to collect and process these communications at a faster pace and on more individuals, can give a great deal of intelligence to draw a map of human relationships (Escudero, 2001). The ability, for instance, to locate and track individuals via mobile phones, which was not available in the past, can provide the Agencies with a large amount of information. It has been suggested that if an Agency combines the traffic data (e.g. who, when, where) of the 100 emails, 50 Internet pages and 20 text messages that an individual might now send or browse in one day, it will have a much richer picture compared to the information it would have acquired from the 10 phone calls and one letter an individual might have sent 25 years ago (Intelligence and Security Committee, 2015).

Current policies acknowledge that the distinction between the content of communications and traffic data provided in 2000 by RIPA, which defined the latter as the basic 'who, where and when' of a communication without defining any other categories of communications, is no longer meaningful as it is now considered to be "technologically obsolete".[51] However, there has not yet been any successful attempt to draw a clear line between the content and traffic data nor has any surveillance law so far taken into account the meaningful information that can be derived from the patterns of a communication and the interaction between individuals. Even though the Data Retention and Investigatory Powers Act (DRIPA) 2014 listed the specific types of data that can be ordered to be retained, there has been a move towards more generality and obscurity in the amendment made by the Counter-Terrorism and Security Act 2015 to cover IP resolution data.[52] The differences are also "incredibly blurred" in the draft Investigatory Powers Bill which, despite its intention to distinguish between the delivery mechanism of the communications data and the content itself, eventually fails to do so, thus possibly resulting in service providers capturing everything just in case something falls under the former category (Science and Technology Committee, 2015). The draft Bill has been criticised for being based upon "a fundamental misconceptualisation"[53] of the significance of communications data, as it treats communications data as less sensitive and makes no reference to the three definitional categories of data identified by the Select Committee which dealt with the bulk collection of data by GCHQ – one that was clearly metadata, one that was clearly content and a grey area termed 'communications data plus'[54] (Science and Technology Committee, 2015). When the Home Secretary was inquired by the Intelligence and Security Committee (ISC) of the Parliament about that distinction, she responded that the nature of communications data is not similar to that of the content and therefore the same process should not be applied, while, in terms of the volume of communications data, the Home Secretary claimed that there has not been such a significant change that should in itself require a different approach. The same position was held by the Agencies which stated that looking at the content of a communication is much more intrusive and that the analysis of the

[51]A term used by the then Chair of the Equality and Human Rights Commission, Baroness O'Neill,during the oral evidence provided to the Intelligence and Security Committee on 14th October 2014.
[52]See Article 2 (3) (b)
[53]Oral evidence of Dr Joss Wright on the technological aspects of the Investigatory Powers Bill provided to the Science and Technology Committee of the Parliament on 10th November 2015.
[54]*See* Intelligence and Security Committee of Parliament, Privacy and Security: A modern and transparent legal framework, 2015, p.53.

communications data helps to reduce intrusion by improving the targeting of more intrusive capabilities. Due to the confusion as to what should be treated as communications data and what as content, particularly in relation to Internet communications and web browsing histories, the ISC has pointed out the need for greater clarity and transparency around the different categories of information in the Bill (Intelligence and Security Committee, 2015).

5.3.3. Excessive powers granted through technological neutrality

A concern often raised with respect to technology neutral policies is that primary legislation in the fields of surveillance and data processing activities does not contain a sufficient level of detail to allow Parliament to scrutinise the proposed measures effectively. In practice, technology neutral laws shift the decisions relating to the technical details to secondary legislation, statutory instruments and their equivalents, which face far less parliamentary scrutiny than primary legislation, while the debates about them are often poorly attended (Whitley, 2013). Nevertheless, the fact that detailed considerations, which may have more severe practical implications, are left to unnamed technocrats and civil servants rather than the elected legislature is inherently problematic and renders those laws potentially subject to abuse. This trend towards shifting crucial decisions to secondary legislation is reflected in the following words: *"We are shifting our legislation to the Executive, effectively, in relying so heavily on secondary legislation, and I think there needs to be greater scrutiny"* (House of Lords Select Committee on the Constitution, 2009). It has been suggested that government's powers should be set out in primary legislation, otherwise there lurks the danger that new powers are granted through technological ambiguity rather than clear debate and process, thus reducing democratic protections and oversight (Escudero-Pascual & Hosein, 2004). In this way, policies set in technology neutral language appear to be to the advantage of policy makers as it is highly possible that *"Agencies will be directed to take advantage of new, technology neutral standards for intelligence gathering"* (Ashcroft, 2001).

An illustrative example of what has been described as "an excessive delegation of powers" is the technology neutral approach adopted in the Identity Cards Act 2006, which introduced a National Identity Scheme that would include the use of biometrics and would be based on a centralised National Identity Register containing the details of all UK citizens (The Identity Projects, 2005). The Identity Cards Act has been strongly criticised for leaving the implementation details of the Scheme to secondary legislation, in spite of the impact on the

relationship between the individual and the State that it would entail (The Identity Projects, 2007). For instance, the statutory purposes of the Act were drawn in broad and flexible terms, such as "for the purpose of securing the efficient and effective provision of public services"[55], while the biometrics to be stored by the Government were not specified, thus leaving some significant details (e.g. whether iris or face biometrics could be taken and recorded) to be added at the procurement stage. Another example of the hazards hidden in a technology neutral surveillance law is the case of the Communications Data Bill 2012. In order to draft future-proofed and flexible legislation able to accommodate future technological innovations of the communications industry, legislators chose to grant "excessively broad powers" to the Secretary of State, who could impose specific requirements regarding data retention with "secret notices" (Anderson, 2015). The Bill was scrutinised by the Parliamentary Joint Committee, which stressed the need for the Bill to be narrowed to cover only the gaps so far identified: "*An undertaking, whether by officials or by ministers, that a power will be used only to a limited extent, is of little value. Once a power is on the statute book, it is available to be used, and also to be misused or abused, at any time in the future"* (House of Lords & House of Commons Joint Committee on Draft Communications Data Bill, 2013).

Similar criticism has been expressed with regard to the data retention powers under DRIPA, which are exercisable by notice from the Secretary of the State to public telecommunications providers. On the grounds that revealing any details about the notices would jeopardise national security, the Government declined to reveal such details even to the court that heard the DRIPA judicial review (Smith, 2015). More recently, critics have targeted the technology neutral language used in the draft Investigatory Powers Bill for not only leading to "*an enormous extension of the overt reach of the state*"[56] but also does not contain any provision for a judicial review procedure to sit behind the warrant for bulk collection of personal data since the judge will only be looking at the investigatory process and not the legitimacy of the reasons for the data collection (Science and Technology Committee, 2015). While scrutinising the Bill, the Intelligence and Security Committee of the Parliament considered it unacceptable to rely on internal policies or Codes of Practice and therefore leaving the safeguards up to the Agencies "as a matter of good practice". In order to highlight the need for clarity and assurance, which the Bill seems to fail to provide, the Committee proposed the

[55] Identity Cards Act 2006, Article 1(4)(e)
[56] Oral evidence of Professor Ross Anderson on the technological aspects of the Investigatory Powers Bill provided to the Science and Technology Committee of the Parliament on 10th November 2015.

enactment of a new Act of the Parliament that should clearly set out the purpose, scale and use of the intrusive powers of the intelligence and security Agencies, while it added that this Act should be distinct from the legal framework governing police and law enforcement authorities as the latter conduct different activities from those conducted by the Agencies (Intelligence and Security Committee, 2015).

5.4. Conclusion

As the above analysis has indicated, given the underappreciated flaws in the arguments in favour of the technology neutral approach, dogmatically embracing the principle of technological neutrality often carries the risk of abandoning legal certainty and clarity with severe implications to the rights to privacy or security. Instead of treating technological neutrality as a default choice or a freestanding principle, it would be wiser to rather consider it as one of the two legislative paths that should be followed depending on the circumstances. In the cases when technology neutral provisions prove to be potentially hazardous, such as in the field of invasive powers where unanticipated activity may inappropriately fall into scope in the future, technology specific laws seem to be the most suitable way to balance the government's need to provide security with the individual's right to privacy. The potential abuse of the Executive's surveillance powers in the absence of proper oversight, combined with the impact of such powers on fundamental rights, may render the need for a clear understanding of the powers authorised more significant than the need for future-proofing. If the law fails to keep up with technological change, in an area of this sensitivity, the Government ought to re-evaluate the new circumstances and accordingly grant new, extended or reduced powers (Anderson, 2015). As stated by Professor Ross Anderson,[57] in an attempt to highlight the dangers involved in drafting future-proofed surveillance laws and the need for those laws to be frequently revised: "*You cannot expect to have a Bill that will last for 25 years*".

At the same time, however, the implementation challenges of embracing the technology specific approach should be also borne in mind. Once legislators decide to create a technology specific law, they will struggle to describe the technology at the proper level of specificity. To make the right decision about how specific a law should be, it is essential that lawmakers gather accurate and complete information about the technologies to be regulated

[57] *Id.*

so as to understand how these technologies operate or how they have been deployed. Striking the right balance between breadth and specificity can be difficult; laws should refer to a certain technology in such detail that allows for the benefits of technological specificity but also generally enough to prevent the need to revisit the law every six months. The second practical difficulty is associated with the clear advantage of technological neutrality over technological specificity, that is, longevity. As technology progresses, technology specific laws expire, quickly or gradually, creating a lack of legal guidance deregulating surveillance, security and privacy protection. This lack of legal certainty during the transition period, however, can be dealt with the existence of background rules that step in to fill the void. Background rules tend to be technology neutral and they apply when technology specific laws expire precisely because they are not closely linked to any particular technology; without technology neutral background rules, the enactment of technology specific laws would be impossible (Escudero-Pascual & Hosein, 2004).

CHAPTER 6: Need for Legal Specificity - The Example of Deep Packet Inspection

6.1. Introduction

Deep Packet Inspection (DPI) is a technique that has been used for several years to maintain the security and integrity of networks, which allows Internet Service Providers (ISPs) to peer into the digital packets that compose a message or transmissions over a network and search for possible threats (Office of the Privacy Commissioner of Canada, 2013). DPI is a type of data processing that looks in detail at the contents of the data being sent and re-routes it accordingly, making sure that a feed of data is supplying content in the right format, or is free of viruses (Geere, 2012). DPI involves inspection points being set up at places on the network, which are not endpoints; DPI technology permits ISPs to look into the content of the packets being sent over the network and act on that information depending on whether it conforms to certain security criteria (Van Eijk & Van Engers, 2010). Potential advantages of DPI Internet monitoring mentioned in the literature include bandwidth management by network providers in order to optimize the network transmission speed and the use for spam filters and virus filters. When a network provider engages in deep packet inspection, it does the equivalent of opening up letters in a postal depot, and reading the contents. Software is used to scan the contents of each packet (and sometimes log it), and then a packet can be re-routed (or dumped entirely) if it passes certain criteria. Those criteria could be the presence of a virus, or just prioritisation of certain types of traffic that are extremely bandwidth-dependent (Mochalski & Schulze, 2009).

DPI is a controversial technique because the legality of its use depends on who is using it, how it is being used and the purpose for which it is being used. Given that not only is the owner of the network infrastructure able to observe the content of the information being sent but can even manipulate the information received over the network, DPI might infringe the principle of confidentiality of communications while it also challenges the open and egalitarian character of the Internet. What is more, the use of DPI technology seems to generate *prima facie* privacy concerns as it has the potential to provide ISPs or other entities with wide access to huge amounts of personal information (Daly, 2010). Without strict

limitations to preserve user privacy, this sort of deep data filtering can significantly impair a user's ability to remain anonymous online (Warwo, 2012). Despite the fact that DPI was originally intended as a means of safeguarding the network by intercepting malicious programmes being sent over the Internet before they reached the end-users, DPI technology has also been put to other uses, such as targeted advertising and government surveillance. DPI allows third parties to draw inferences with respect to one's interests, purchasing habits or other activities and thus targeted advertising is based on one's behaviour while browsing the Internet (Bendrath, 2009). In addition, DPI can be used for purposes of criminal investigations, and thus enables governments to monitor communication by capturing and recording packets in the name of national security (Anderson, 2007). A key notion for the legal use of DPI is transparency, which constitutes a requirement for the legitimate processing of personal data according to the DPD. Therefore, if ISPs inform their customers (the 'data subjects') of the use of DPI technology, and seek their consent on an opt-in basis before their data is processed, in this case DPI might provide some privacy guarantees; however, as far as sensitive personal data is concerned, the use of DPI still remains problematic, even if ISPs meet the above transparency requirements (Daly, 2010). The problematic nature of DPI techniques in terms of the discretion granted to ISPs when applying security measures is the focal point of this chapter, which examines the implications of such techniques to individual's privacy. To this end, the chapter starts with explaining the way DPI works providing technical details that shed light to the intrusive character of DPI techniques and the hazards involved in its implementation if used inappropriately. After presenting examples indicative of the problematic nature of DPI, the chapter proceeds with the legal safeguards that should be taken into consideration when making laws concerning techniques that have the potential to impact on individuals' fundamental rights and violate the basic principles of Internet governance.

6.2. Deep Packet Inspection: Overview

6.2.1. Packet header and packet payload

The demarcation line between the packet header and the packet payload is a fundamental aspect of the definition of DPI. To begin with, when data are transferred over the Internet, they are broken down into multiple packets – units of binary data capable of being routed

through a computer network – and then reassembled to the original data chunk once they reach their destination; each unit transmitted includes both a header and a payload. Internet packets do not have only a single header and payload; instead, there is a packet header and payload at each layer of the multi-layered Internet architecture that can be found in each network-connected host (Mochalski & Schulze, 2009). The header contains transmission-related instructions about the data carried by the packet, which tell routers how to handle and forward the packet along to its destination. Header information often includes the packet's total length, originating address (where the packet came from), destination address (where the packet is going), sequence number (which packet this is in a sequence of packets), protocol (what type of packet is being transmitted, e.g. e-mail, web page, stream video). The payload, also called the body or data of a packet, is the cargo of a data transmission. The packet payload contains information about the actual content the packet carries, what application is sending the data, whether the contents are themselves encrypted (Fuchs, 2013b). More specifically, the payload consists of application data (e-mail text, URL[58], website content, chat message, video content, image content etc.) as well as application header (application programme version, e-mail address sender/receiver etc.). In a network packet, headers are appended to the payload for transport and then discarded at their destination.

In order to better understand the distinction between the packet header and packet payload, a brief analysis of the structure of the most widely used communication standard, the TCP/IP, is necessary.[59]The TCP/IP[60] networking model was developed over thirty years ago to standardise how network devices communicate in an interoperable fashion (Tanenbaum & Wetherall, 2010). The TCP/IP protocol suite comprises four layers – link, network, transport and application – plus the physical layer over which the model runs, each of which implements a subset of functions necessary for end-to-end data transmission. The *physical layer* defines the actual media over which the data are being transmitted; the *link layer* formats the packet so that it can be sent from its point of origin to its destination, or where necessary to the next router towards the destination; the *network layer* is responsible for the packet's addressing and routing; the *transport layer* organises the data transmission process in several sequential steps by segmenting data from upper levels and reassembling the data flow into smaller units, and (in the case of TCP) establishes a connection between the

[58]The Code of Practice for the Acquisition and Disclosure of Communications Data (2003) provides that the part before the first slash in a website address is communications data, and what comes after the first slash is content; therefore, full URLs are regarded as content.
[59] In order for the data to be transmitted within and across the network, communication standards are needed.
[60] TCP stands for Transmission Control Protocol and IP for Internet Protocol.

packets' sender(s) and recipient(s); the *application layer* interacts with the software applications that are making a data request. The first three layers - link, network and transport - each adds a header to the packet, whereas the application layer manages the packet payload (Cooper, 2011). The transport layer lies between the network layer and the application layer, and thus a transport header exists between the IP packet header and the packet payload. The transport header indicates the source and destination applications at the communicating endpoints, e.g. a web browser and a web server, and these are identified by 'port numbers', with certain applications generally (but not always) using well-known port numbers, such as port 80 for unencrypted web (HTTP[61] protocol) traffic (Bendrath & Mueller, 2011). One form of metadata for network traffic is the so-called '5-tuple' which is commonly used in various networking contexts to identify specific application flows. A 5-tuple refers to a set of values that comprise a TCP/IP connection and is the combination of source and destination IP addresses and ports, together with the protocol in use (Chen et al., 2010).

6.2.2. The different types of packet inspection

Depending on the TCP/IP layers that packet inspection technologies can analyse, there are three 'classes' of these technologies that are used in networking environments; shallow, medium and deep in the sense of whole packet inspection (Porter, 2010). The TCP/IP model can be used to express the extent of information that inspection technologies can derive from packets; the closer such a technology comes to examining the application layer part of the payload, and the further it looks into the payload, the more information it can learn about the packet passing through the inspection device (Parsons, 2008). Shallow Packet Inspection (SPI) examines the packet's IP and transport header information to decide whether to allow the packet to pass. It might make a decision to only allow traffic to port 80 or 443 (the well-known ports for plain or encrypted web traffic) to certain IP addresses. Or it might check the destination IP address against a blacklist (e.g. of known malware command and control servers); if the IP address is on the blacklist, the packet is not delivered. With SPI it is not possible to peer inside a packet's payload to survey the packet's content (Daly, 2010). If an SPI system notes information about a packet, and uses that to influence decisions on future packets seen, it is known as a 'stateful packet inspection' system; an example is that such a system may let UDP traffic out, and only allow UDP traffic back in to its network that match the traffic sent out. The second class of packet inspection technologies – Medium Packet

[61] HTTP stands for Hypertext Transfer Protocol.

Inspection (MPI) - involves the use of application proxies, devices that act as intermediaries between end-users' computers and Internet gateways, through which all the traffic passes. An example might be a web cache or proxy, used by an ISP or a site such as a university campus network. Application proxies examine packet headers and a small amount of payload against parse-lists for particular representations; every parse-list contains a set of representations. MPI devices decide whether a specific packet-type is permissible or not according to its data format type and its associated location on the Internet. Due to their ability to read the application commands located within the application layer as well the file formats in the presentation layer, MPI devices can prevent users from receiving specific types of files or prioritize specific files over others (Parsons, 2008).

Whereas MPI devices have very limited application awareness, DPI devices are designed to allow ISPs using them to precisely identify the origin and – more importantly – the content of each packet that passes through their networks. DPI devices can examine all the headers as well as the whole content of the messages due to their ability to look at every layer of the TCP/IP model (Fuchs, 2013a; Office of the Privacy Commissioner of Canada, 2009). Whole packet inspection technologies enable ISPs to gain greater control over every facet of their network operations (Cooper, 2011). Many of the functions provided by DPI technology have been available before; what differentiates DPI from other technologies are the following unique characteristics. First, in addition to inspecting the packet header, not only can DPI use any part of the packet for detection, but it also looks for patterns across multiple packets (Mochalski & Schulze, 2009). Another key feature of DPI is that, unlike several applications that scan digital content stored on servers or computers, DPI scans information in motion, not information at rest. The scanning of data packets takes place in real time; that is why DPI is much faster than its predecessors (Artan & Chao, 2007). What is more, DPI equipment allows network operators to make decisions that involve more than merely where to forward the packet; network operators monitor the content of data packets in real-time and make decisions about how to handle them (Parsons, 2012).

6.2.3. DPI capabilities

DPI technology is widely recognised as a powerful tool used for inspecting, deterring and deflecting malicious attacks over the network (Abuhmed et al., 2008). The exact meaning of DPI, however, is quite vague since it is not a standardised technology; the way it is deployed

is not usually disclosed as it is considered the proprietary information of each DPI vendor (Moon & Kim, 2014). DPI allows for visibility into the application layer as DPI devices can look at the actual data traversing the network rather than just keeping track of connection information, and thus understand what the application is on that network flow[62] (whether it is a video, file transfer, low-bandwidth gaming, text communication etc.). By stripping off the headers, DPI devices can use the resulting payload to extract information from traffic that varies by application type; IP addresses and URLs from HTTP traffic, SIP numbers from VoIP calls, file names of P2P files, and chat channels for instant messages.

DPI has become an essential tool for network operators in their effort to obtain an in-depth knowledge of the underlying traffic composition and dynamics and thus be able to intervene in the management of that traffic. Specific applications can be passive aiming solely at giving the operator greater visibility, which is crucial for investment or pricing decisions, capacity planning etc. (Mueller, 2011). When active elements are added to passive applications, DPI can contribute to bandwidth regulation or congestion response; DPI enables operators to discriminate among different types of traffic in order to provide satisfactory quality of service, or to throttle down excessive traffic (Renals & Jacoby, 2009). When the recognition capability of DPI technology is combined with the manipulation capability, ISPs can prioritise (or de-prioritise) specific protocols, services or users. This market opportunity created by deeper and more granular traffic inspection allows ISPs to differentiate the online experience of individual users not only by bandwidth tiers, but on application and content bases as well; ISPs may apply different charging policies, traffic shaping, or offer quality of service guarantees to selected users or applications (Antonello et al., 2012). An example of deploying DPI technology for such purposes is given by the Cisco Service Control Engine, which segments its customer base in order to provide individualised services to differing demographic ‘clusters’ by combining DPI with specific subscriber identification (Frieden, 2008). DPI has been described as “a potentially disruptive technology” (Bendrath & Mueller, 2011) as DPI used by ISPs can result in a violation of longstanding standards, such as net neutrality, and create a tiered Internet, which disadvantages certain users and application types, and is controlled by large companies (Fuchs, 2013a).

[62] When an application running on a user’s computer initiates a communication with another computer on the Internet, the sequence of packets exchanged between the two computers is known as a *flow*. Each individual user and many applications can sustain multiple flows simultaneously.

6.2.4. Depth & breadth of Deep Packet Inspection

6.2.4.1. The depth of Deep Packet Inspection

There has been some controversy about how ‘deep’ the inspection of packets should be for it to qualify as DPI. Although the most widely adopted definition of DPI makes a distinction between packet headers and packet payload, the conception of ‘deep’ in terms of inspecting network traffic often varies. The strictest conception of ‘deep’ describes DPI as the use of any data other than the destination IP addresses (Bowman, 2009), while a slightly more expansive conception draws a line between IP headers and the rest of the packet (Reed, 2008). As a matter of fact, the depth of DPI – the extent to which DPI tools delve into individual packets - measures how much of each packet is subjected to the abovementioned pattern matching process (Cooper, 2007). The choice of ISPs in terms of how deep the packet inspection should depend on each DPI application - some DPI uses require the inspection of only some portions of the packet payload and some other the inspection of the entire payload - as well as the perceived accuracy of each alternative.

For instance, when DPI is used to identify P2P protocols as part of congestion management or prioritised service offering, it may be sufficient to inspect only a part of the payload since the P2P protocol names are always visible at the same location inside the packet (Mochalski & Schulze, 2009). In case an ISP wants to generate usage data about the amount of email traffic on its network, the ISP could choose to inspect only port numbers to identify common email ports in case the application uses the commonly used port, or it could inspect entire payloads to look for well-known email protocol signatures. However, other uses such as blocking viruses, filtering objectionable content, or behavioural advertising, necessitate the inspection of the entire packet payload in order to determine if the network traffic matches a virus signature or a content fingerprint, or contains information that indicates a particular user’s interest. An example of DPI’s depth in terms of network security is the product of DPI vendor Dell SonicWall, Dell SonicWALL’s Reassembly-Free Deep Packet Inspection. As mentioned on the company’s website, DPI engine is capable of determining exactly what applications are being used and who is using them by scanning “every byte of every packet of all network traffic”; in order to protect networks against the sophisticated attacks that target

application vulnerabilities, “all downloaded, emailed and compressed files are examined at the application layer”.[63]

6.2.4.2. The breadth of Deep Packet Inspection

Another concept worth mentioning is the ‘breadth’ of packet inspection, which refers to the number of packets inspected in a flow as well as the total number of subscribers whose packets pass through a DPI engine (Cooper, 2007). As far as the number of packets is concerned, conducting a pattern match on only the first few packets in a flow may provide sufficient information to the ISPs; for instance, when filtering content of particular types or when trying to spot the start-up of a P2P session, the ISP can inspect only the first packet and then apply its rules to the entire flow without further inspecting the rest of the packets. Similarly, some networks experience congestion much more frequently in one direction (upstream or downstream); ISPs can limit their DPI use to the relevant direction only thus limiting the number of packets that get inspected. Two examples indicative of the breadth-related design decisions are those of the behavioural advertising companies Phorm and NebuAd; although DPI was deployed for the same purpose, the breadth of packets involved in Phorm’s system was far greater than that of the NebuAd system. Whereas the system used by NebuAd firm required inspection only in one direction (it captured only URLs and search terms), Phorm inspected traffic flowing in both directions, collecting data about both the user’s website request (URLs and search terms) and the website content sent to the user (Clayton, 2008; McCullagh, 2008).

The purpose for which DPI technology is deployed also defines the number of subscribers whose packets are inspected; ISPs can deploy DPI throughout the network so that all subscribers’ packets are subjected to inspection, or they may selectively deploy it on certain groups of subscribers or at certain network nodes. A sample of network nodes may be sufficient to give an ISP an indication of how many subscribers are using the network or to diagnose a specific problem,[64] but behavioural advertising or content filtering may require the monitoring of every single subscriber. In addition to DPI application, DPI’s breadth also depends on whether a DPI system proactively monitors traffic or only in response to signals. For proactive spam management, for example, an ISP may need to scan every email message

[63]<http://www.sonicwall.com/us/en/products/Deep-Packet-Inspection.html>.
[64]This is the case of Cleanfeed (*see*3.3. DPI private experiments) where only traffic to the “blacklisted” IPs is subject to DPI to see if the precise URL is deemed blocked.

to determine whether each message matches the predefined spam signatures,[65] whereas ISPs employing the troubleshooting approach might wait for users or other ISPs to report the problem and then use DPI for investigation purposes. There are some cases, however, where the only reasonable case is the proactive deployment of DPI, such as in behavioural advertising or prioritised service offerings, because it is the continual monitoring of the network traffic that allows the ISP to offer the DPI-based service.

6.2.5. The role of Deep Packet Inspection in network security

The major drivers that are shaping the new security landscape – increasing complexity of networks, increasing sophistication of applications and attacks etc. – have led to a radical evolution in the nature and requirements of network security the last few years. Noteworthy is the shift from so-called 'network-level' threats such as connection-oriented intrusions, to dynamic content-based threats such as viruses, worms, spyware, which target applications and application layer protocols rather than the networks they are transported on (e-Soft, 2013). These modern threats can spread quickly and are designed to bypass traditional firewalls; therefore, sophisticated levels of intelligence are required to detect and stop them. DPI adds another layer of intelligence to the existing firewall capabilities as it is produced by the convergence of traditional approaches in network security (Ramos, 2009).

Network security has been, in fact, the earliest driver of the development of DPI technology; DPI can protect the network from attacks by monitoring, identifying and throttling traffic at all layers of the TCP/IP model. DPI technology was first developed for intrusion detection and intrusion prevention systems. By combining malware recognition capabilities with packet capture and analysis techniques, DPI engines allow network operators "to detect and intercept recognised forms of threats before they reach their customers, such as keystroke loggers, bot infections, abnormal quantities of mail, command and control instructions from bot herders, or communications to servers known to be associated with botnets" (Mueller, 2011). Not only are DPI engines able to identify threats but they can also respond to the detected threats by implementing measures to prevent the attack from succeeding, usually by terminating connections. More specifically, DPI technology allows a security application to peer deep into the content of a data stream; certain patterns or signatures need to be found in the packet data in order for DPI to detect an attack on the network or other malicious behaviour.

[65] The scanning of email messages does not have to be real-time; it can take place on the server, not in transit.

As a matter of fact, the predecessors of DPI restricted ISPs' view to network-level details, at the network and transport level, where spam and viruses are hard to distinguish from other email messages or web surfing behaviour. At this level, ISPs can guess from the port number, e.g. 80/http, but the applications may run on any port; therefore, correctly identifying an application requires inspecting the payload itself. DPI tools can identify, for instance, viruses, by comparing files crossing a network to a database of known viruses; spam, by analysing the words used; and intruders, by looking at the commands they send (Anderson, 2007). DPI is considered a vital component for network security as it provides ISPs with application awareness (DPI devices are capable of understanding both the protocols and the actual applications that rely on those protocols), which enables ISPs to intelligently read and analyse the huge volumes of data transferred over their networks. As a result, ISPs can quickly gain accurate application visibility and control over security threats and thus take prompt action to prevent network attacks. What is more, DPI devices are designed to scale in large networking environments and thus determine what programs generate packets, in real-time, for hundreds of thousands of transactions each second.

6.3. Deep Packet Inspection Limitations and Dark Sides

6.3.1. Deep Packet Inspection limitations

As analysed above, DPI technology is of vital importance for network providers as many critical network services rely on the inspection of the packet content, which gives a broader picture than only looking at the structured information found in the packet header (Antonello et al., 2012). Nonetheless, despite its huge capabilities DPI is not without limitations. DPI capabilities depend on the ability to define algorithms or codes that accurately capture the features of what is sought in the traffic stream (Bendrath & Mueller, 2011). Given the huge number of applications represented by signatures, the patterns a DPI engine searches for must be constantly updated and expanded to deal with new or changing phenomena. The large number of signatures creates a serious scalability constraint on DPI implementation since DPI devices may not scale well with link speeds possibly resulting in slowing down the Internet connection (Kim at al., 2008). Moreover, DPI systems may be accurate but they are also resource-intensive because the process of real-time recognition of patterns in high speed networks requires highly sophisticated algorithms thus imposing heavier management

burdens on network operators. It is becoming increasingly challenging for DPI to be performed in real-time on high-speed Internet backbone networks, where speeds at the time of writing are commonly now 100Gbit/s. It is also useful to note here that effective DPI is only possible when the payload is plain text; where encryption is used the ISP would then need to be able to break that encryption in order to inspect the payload. Recently, both Facebook (due to security concerns over account hijacking) and Google Mail (post Snowden revelations) have made traffic to their systems encrypted by default. In cases where IPsec[66] is used, which may include virtual private networks (VPNs), other layer header information may also be unavailable to a third party monitoring system, the specifics depending on whether end to end or tunnel mode IPsec is used.

6.3.2. Deep Packet Inspection dark sides

6.3.2.1. DPI implications on privacy

DPI is a highly controversial technology which has a lot of societal implications that need to be carefully considered. The fact that its use allows network operators to look at the content of communications may be crucial to bandwidth management by network providers in order to optimize the network transmission speed and the use for spam filters and virus filters, but it also raises concerns about the possible misuse of such a powerful technology that is hard to control (Fuchs, 2013a). The most commonly mentioned fear is that DPI leads to an "unreasonable invasion of an individual's privacy" (Office of the Privacy Commissioner of Canada, 2009) because there is no limit to the data one can extract from the payload, as long as they can understand the payload (Gallagher, 2012). Not only does DPI technology invade the privacy of one's communications by providing access to the content itself, but it also enables third parties to draw inferences about users' lives and based on them they can further construct "vast network social maps" (Parsons, 2008). DPI enables network operators to identify not only the tools used to communicate, the IP addresses that communications data are being transmitted to or the number of intended recipients, but it also makes it possible to identify the perceived relationships between individuals transmitting data to one another. In theory, someone in control of a DPI device could read a user's email messages, see every

[66] IPsec stands for Internet Protocol Security. This protocol suite is used to secure IP communications and relies upon authentication and encryption techniques.

web page exactly as they saw it, and easily listen to their instant messaging conversations (Jason, 2011). At this point, it is worth wondering whether privacy protection should be a default choice guaranteed by the ISPs or a non-default option that can only be achieved by certain actions on behalf of the users. Even though it is likely that users allow the monitoring of their online activities in exchange for other things they value more, it must be a deliberate exchange. Not only should the users have that choice in an explicit way, but DPI-based network functions must be clearly explained in advance since a typical user cannot be expected to understand the mechanics of the wen sufficiently. For instance, a way to mitigate the risk of privacy intrusion is to implement DPI functions not automatically in a network, but to give the user the possibility to opt-in to network-wide scanning for certain data types (e.g. emails) for limited purposes such as detecting spam mail and viruses.

6.3.2.2. DPI implications on net neutrality

DPI violates fundamental qualities of the Internet, such as net neutrality, and it has the potential to alter Internet operations and governance. Network's ability to discriminate between applications and senders results in the creation of different tiers of online service which entails information inequality (Lessig & McChesney, 2006), while it undermines consumers' trust that their online communication will be delivered without interference (Cooper, 2011). What is more, it renders ISPs gatekeepers of cyberspace, thus increasing the likelihood of imposing intermediary responsibility on them. The question that arises is whether or not ISPs, whose role is to provide access to the Internet, should be entitled to leverage their position as providers of basic transport for use in other ways by serving their market on a discriminatory basis. A tiered Internet monitored with the help of DPI could also result in users turning to encryption to avoid the monitoring of their data, which in turn leads to other forms of inequality as only the technically skilled users will be able to protect their data (Anderson, 2007). The heavy use of encryption would also render the congestion-reduction purpose of DPI ineffective since it could slow down the network connection speed (Riley & Scott, 2009). However, can users be expected to use encryption for all or large parts of their Internet use or should ISPs be in charge of protecting the confidentiality of their users' data?

6.3.2.3. DPI and surveillance

The advent of DPI technologies revolutionised network surveillance over the last decade and restructured the range of surveillance Internet users are subject to because DPI capabilities (e.g. grabbing information from network traffic in real time) can be used for more wide-ranging monitoring of the users. The fact that DPI technologies allow for monitoring of not only the metadata of Internet communications (sender, recipient, type of data etc.) but also the sent content render them useful tools in the hands of state and commercial actors that want to monitor citizens' and consumers' behaviour.[67] It has been argued that the implementation of DPI had as a result the emergence of "a new mode of governmentality" (Fuchs, 2013b), in which the state and security industry interests interact. DPI technologies, which were initially developed for the purpose of fighting cyber threats, could end up being a lucrative business for companies that produce and sell monitoring technologies to the state actors. As a matter of fact, the recent Snowden revelations showed the level of collaboration between the telecommunications companies and the surveillance agencies (McAskill, 2014).

With the help of DPI, ISPs hold the power to potentially build Internet surveillance system, in which most or the entire Internet usage will be monitored. A typical explanation given by security companies for selling Internet surveillance technologies is associated with the idea that criminals use the Internet; therefore Internet surveillance is necessary to prevent crime terrorism. Based on the fact that a DPI application can make anything that happens on a network visible and recordable to governments, national laws typically require communication service providers to provide some kind of eavesdropping to government law enforcement or public security agency. In liberal democracies, such capabilities are usually controlled by laws which regulate, for example, what kind of persons are subject to surveillance, what information can be collected, how long it can be stored, what is admissible evidence for a prosecution etc.; this practice is called 'lawful interception'. Law enforcement could set up an automated system that issues demands to the ISP(s) in question, and then establish procedures to gain access to subscriber information. This notion of automated mass surveillance for law enforcement purposes leads to the question of what drivers should motivate the technology and what the long-term effects of using DPI for detecting criminal behaviour would be. When the government in question is authoritarian, technologies such as

[67] See Sophie Stalla-Bourdillon, Evangelia Papadaki and Tim Chown, 2014, From Porn to Cybersecurity Passing by Copyright: How Mass Surveillance Technologies Are Gaining Legitimacy… The case of Deep Packet Inspection Technologies, *Computer Law & Security Review* 30 (2014): 670-686.

DPI can be used for arbitrary, unrestricted surveillance of the contents of user communications. As stated by Ross Anderson, professor in security engineering at the University of Cambridge Computer Laboratory, DPI technology is used to monitor citizens' communications in repressive regimes which often have laws making repression legal (Out-Law.com, 2012). There is the risk that the processing and analysis of sensitive content data results in the political repression or social discrimination of certain groups.

According to David Lyon's definition, surveillance is "the focused, systematic and routine attention to personal details for purposes of influence, management, protection or detection" that it is "deliberate and depends on certain protocols and techniques" (Lyon, 2008). In fact, potential uses of DPI technologies, which aim to influence or manage the traffic flow by either altering the data packets or limiting the online actions that can be performed, appear to fall under this definition. The same applies to the cases where DPI is deployed to secure the networks; the need to perform regular packet analysis in order to detect heuristic matches and anomalies also corresponds with Lyon's definitional elements of 'protection' and 'detection'. However, there exist two more factors that should be taken into account before jumping to conclusions regarding the connection between DPI and surveillance. First, it is important to examine whether DPI devices are *generally* used for all the above purposes (Parsons, 2008). Second, surveillance needs to be distinguished from search, two terms that can easily be confused. Whereas search entails the specific targeting of an information type, surveillance extends beyond search as it involves the inspection of each packet that passes along the network that might accidentally capture information beyond that sought (Solove, 2008). For example, while surveillance may identify how many popular VoIP applications are on an ISP's network, a targeted search will reveal precisely which applications are used by a specific individual. Similar is the distinction between strategic and individual monitoring as described by the European Court of Human Rights in Weber and Saravia v Germany[68], where the court defined strategic monitoring as the act of "collecting information by intercepting telecommunications in order to identify and avert serious dangers facing the Federal Republic of Germany, such as an armed attack on its territory or the commission of international terrorist attacks and certain other serious offences". In contrast, individual monitoring is described by the court as the interception of telecommunications that involve "specific persons suspected of having committed or planning to commit serious crimes, in order to prevent or investigate those offences."

[68] Weber and Saravia v Germany (2006) 1173 E.C.H.R. at [4].

6.3.3. DPI private experiments

6.3.3.1. Cleanfeed

In order to deal with illegal child sexual abuse content available on the Web, the UK has adopted an industry-led approach. Since 1996, the Internet Watch Foundation (IWF)– a private body funded by the Internet industry and the EU – has acted as a hotline receiving public complaints and, in case these complaints are grounded, forwarding them to the police as well as asking the UK-based hosting providers to have that material removed.[69] In spite of the remarkable success of this approach, it was not effective when the illegal material was hosted abroad (McIntyre, 2013). To address this problem, British Telecomm (BT) developed a technical system, called Cleanfeed, aiming specifically at blocking access to child abuse images hosted outside of the UK jurisdiction. The stated purpose of Cleanfeed was to prevent Internet users from accessing, either accidentally or by design, illegal child abuse images (Clayton, 2005).

The Cleanfeed system was designed to be a low cost, but highly accurate, method of blocking unacceptable content; this filtering technology is a hybrid system, which combines the redirection of traffic (IP address re-routing) and DPI-based URL blocking, and operates as a two-stage mechanism to filter specific Internet traffic (House of Commons Culture, Media and Sport Committee, 2007-2008). The first stage is based on the examination of the IP address and the destination port of the packets travelling across the network against the IWF database. In case the traffic is suspicious, then it is redirected to the second stage of filtering, which is implemented as web proxy that understands HTTP requests. When the request matches an item from the IWF blacklist containing URLs of child sexual abuse content, a *404 response* is returned; 404 response is a HTTP standard response code indicating that the server could not find what the client requested (Clayton, 2005). At the time of its implementation, the Cleanfeed technology was regarded as a substantial step forward over the two main schemes of filtering then in use; in contrast with the IP address blocking, the use of web proxies, which were as selective as necessary, prevented the danger of over-blocking, while it ensured that only web traffic – no other protocols – would be affected by avoiding the second existing scheme, DNS poisoning (Clayton, 2005).

[69] Internet Watch Foundation <https://www.iwf.org.uk/>.

Despite the fact that BT designed the Cleanfeed system in such a way as to be used only for the protection of its customers – not for prosecution purposes – by avoiding logging data on users (Hutty, 2004), its implementation has been described as "the first example of mass censorship on the Web attempted in a Western democracy" (Bright, 2004). What prompted public criticism was the power given to a private body to take censorship decisions for the UK Internet users by determining whether the content is potentially illegal, with limited procedural safeguards and no oversight (Davies, 2009). An example of the implications that a filtering system such as Cleanfeed could have is the Wikipedia incident (Clayton, 2009). In December 2008, the URL of a Wikipedia article was added to the IWF list because of containing a pornographic image of a child, which was the cover image of a Scorpions album being legally sold since 1976. As a result, UK traffic for Wikipedia was redirected via the Cleanfeed servers; instead of blocking access to the specific URL of the offending image, access to the entire page of the band was blocked. After reviewing the situation, the IWF removed the listing four days later (Naughton, 2008).

6.3.3.2. Phorm

In 2008, a company named Phorm, whose aim was to help advertisers better target consumers by monitoring their web browsing habits, raised much ire from privacy campaigners. Phorm was in partnership with three of the UK's biggest ISPs – BT, Virgin Media and Talk Talk – which were planning to use the Phorm service, a patent-pending technology that delivered personalised content and advertising to ISP customers (Stevens, 2008). As it was revealed, BT had already carried out secret small-scale trials in 2006 and 2007, where, in both cases, the users were unaware of the tests and thus had not actively given their consent. Phorm Service deployed DPI-based advertising technique, which involved equipment installed in an ISP's network that intercepted all web traffic passing along every customer's broadband connection, and scanned through it for keywords that could be used to deliver targeted advertising – advertising that reflects customers' communication flows. User's IP addresses were said not to be stored (Wray, 2009).

The main criticism of Phorm, as it was planned to be implemented in the UK, focused on the fact that it would have been an opt-out service, which meant that no prior consent would have been required for the users' data to be monitored (McSaty, 2011). Some even stated that even if users decided to opt out of the service, they would still continue to have their browsing

histories stored by Phorm (Heron, 2009). Fearing that such a technology could set a worrying precedent, that intercepting technologies would be perfectly acceptable for commercial reasons, many government officials questioned the legality of an opt-out programme; in April 2008, the Information Commissioner's Office stated that Phorm's system would only be legal under UK law as an opt-in service. The Phorm scandal eventually triggered the European Commission into launching an Infringement Proceeding against the UK, in April 2009, for failing to fully implement the DPD and the e-Privacy Directive (European Commission, 2009).

6.3.3.3. Detica CView

A few months later, in an attempt to measure the level of music copyright on its network via peer-to-peer protocols, Virgin Media UK planned to deploy Detica CView technology on a trial basis beginning at the end of 2009, which would have involved monitoring 40 per cent of its customers without their knowledge or prior consent (ISP Review, 2010b). CView is a DPI product based on the same technology that powered the controversial Phorm's advertising system; it identifies peer-to-peer packets and then it looks at the actual content of those packets including application data in order to determine whether the copyrighted work exchanged is licensed or not, based on data provided by the record industry (Williams, 2009b). CView was designed to look for three types of file-sharing traffic - eDonkey, Gnutella and BitTorrent (Williams, 2009a). As the company's aim was to establish an 'index' of copyright infringements – not to keep records on individual customers – Virgin emphasized that data on the level of copyright infringement would be aggregated and anonymised (ISP Review, 2010a). Indeed, once CView identified an eDonkey, Gnutella or BitTorrent session, it would strip out the IP address of the user from each packet replacing it with a randomly-generated unique identifier and pulling out an 'acoustic fingerprint'. It would then send the processed material on to a central server to be matched against a database of acoustic fingerprints of copyright songs provided by record companies (Williams, 2009a). However, despite the fact that Virgin claimed that processed data would be anonymised, privacy campaigners protested against the implementation of such technology arguing that using that technology to identify those using torrents or blocking the content would only require "a slight tweak to the software" (Nickson, 2009). Several events following the announcements of Virgin made the company reconsider the use of the CView system and put the trial on hold (ISP Review, 2010b).

6.4. DPI and the Law

6.4.1. Data protection and privacy

Depending on the circumstances of each case and on the type of analysis performed, the processing of personal data may be highly intrusive for an individual's privacy (European Data Protection Supervisor, 2011). The data protection implications derived from inspection techniques vary substantially depending on which technique is used. For ISP practices entailing the inspection of individuals' communications, Recital 28 of the DPD amending the Universal Service and e-Privacy Directives highlights that "depending on the technology used and the type of limitation, such limitation may require user consent under the e-Privacy Directive". Thus, Recital 28 recalls the need for consent pursuant to Article 5(1) of the e-Privacy Directive for any limitations based on monitoring of communications. In the UK the use of DPI is legally problematic as Section 8 of the European Convention on Human Rights does not allow warrantless mass surveillance (Out-Law.com, 2012). In Europe, privacy is protected by Article 8 of the European Convention on Human Rights (ECHR), an obligation primarily pertaining to contracting States. It is also protected by the DPD, which regulates the processing of personal data and concerns to organisations in the private as well as public sector.

6.4.1.1. Legal grounds for safeguarding the security of the service

The EU data protection law seems to legitimise the processing of personal data through the means of DPI techniques; the DPD, combined with the e-Privacy Directive, opens the door to a large amount of processing for the purpose of, amongst other, ensuring traffic management and/or network security. Under Article 7 of the DPD, the processing becomes legitimate if it is necessary for the legitimate interests pursued by the data controller, while the e-Privacy Directive expressly identifies several types of legitimate interests, not all of which require the prior consent of the data subject. As far as network security is concerned, pursuant to Article 4 of the e-Privacy Directive, an ISP is under a general obligation to take appropriate measures to safeguard security of its services. The practice of filtering viruses may involve the processing of IP headers and IP payload. Taking into account that Article 4 of the e-Privacy Directive requires ISPs to ensure the security of the network, this provision legitimises

inspection techniques based on IP headers and content that aim strictly to achieve such purpose. In practice, this means that, within the limits set forth by the proportionality principle, ISPs may engage in monitoring and filtering of communications data to fight viruses and overall ensure the security of the network.[70] In the same line, Recital 39 of the GDPR also recognises the legitimising effect of ensuring network and information security. Recital 39 states that "the processing of data to the extent strictly necessary for the purposes of ensuring network and information security […] constitutes a legitimate interest of the concerned data controller"; some examples given include "preventing unauthorised access to electronic communications networks and malicious code distribution and stopping 'denial of service' attacks and damage to computer and electronic systems".

6.4.1.2. Obstacles to the deployment of DPI

6.4.1.2.1. Sensitive data

An obstacle to the deployment of DPI could derive from Article 8 of the DPD which prohibits the processing of sensitive data. Under Article 8, "[m]ember States shall prohibit the processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life". However, it is still possible to process such data if the data subject has given his explicit consent unless the Member State, whose law is applicable, does not recognise the legitimising effect of the explicit consent of the data subject in these circumstances. It is noteworthy that even if the collection of the communications content, which contains sensitive data, could be prevented, the same does not apply to the metadata; therefore, a significant amount of data would still be available for processing.

6.4.1.2.2. Confidentiality of communications

Filtering, blocking and inspecting network traffic raises important questions, often overlooked or side-lined, regarding the confidentiality of communications and the respect for the privacy of individuals and their personal data when they use the Internet (European Data Protection Supervisor, 2011). By inspecting communications data, ISPs may breach the

[70] Article 29 Working Party's Opinion 2/2006 on privacy issues related to the provision of email screening services, adopted on 21 February 2006 (WP 118). In this Opinion the Working Party considers that using filters for the purpose of Article 4 can be compatible with Article 5 of the e-Privacy Directive.

confidentiality of communications, which is a fundamental right guaranteed by Article 8 of the European Convention for the Protection of Human Rights and Fundamental Freedoms (the 'ECHR') and Article 7 and 8 of the Charter of Fundamental Rights of the European Union (the 'Charter'). Confidentiality is further protected in secondary EU legislation, namely Article 5 of the e-Privacy Directive. Article 5 provides that "[m]ember States shall ensure the confidentiality of communications and the related traffic data by means of a public communications network and publicly available electronic communications services, through national legislation. In particular, they shall prohibit listening, tapping, storage or other kinds of interception or surveillance of communications and the related traffic data by persons other than users, without the consent of the users concerned, except when legally authorised to do so in accordance with Article 15(1)". According to Article 15(1) of the e-Privacy Directive, which refers to Article 13 of the DPD, the principle of confidentiality can be restricted for certain purposes (i.e. "the prevention, investigation, detection and prosecution of criminal offences"). Likewise, one could argue that ISPs should be allowed to implement DPI practices and thus set aside the principle of confidentiality in the name of network security.

6.4.2. Net neutrality

Net neutrality refers to an ongoing debate over whether ISPs should be allowed to limit, filter, or block Internet access or otherwise affect its performance. The concept of net neutrality builds on the view that information on the Internet should be transmitted impartially, without regard to content, destination or source, and that users should be able to decide what applications, services and hardware they want to use. This means that ISPs cannot, at their own choice, prioritise or slow down access to certain applications or services such as Peer to Peer ('P2P'), etc. When ISPs inspect communication data in order to differentiate each communication flow and to apply specific policies, which may be unfavourable to individuals, the implications are significant (European Data Protection Supervisor, 2011).

It seems that at the EU level there is a broad aspiration to an open Internet; in spite of the lack of explicit EU rules on net neutrality, certain requirements are set to promote this concept. More specifically, article 22(3) of the Universal Service Directive empowers national regulatory authorities to impose, if necessary, minimum quality of service requirements on ISPs in order to prevent the degradation of services and the hindering or slowing down of

traffic over public networks. Pursuant to article 8(4) of the Framework Directive, national regulatory authorities should take all necessary measures to ensure that users can access and distribute information or run applications and services of their choice. Nonetheless, this policy objective, which applies to the network as a whole, is not directly linked to prohibitions or obligations on individual ISPs. In other words, an ISP could engage in traffic management policies, which may exclude access to certain applications, provided that end-users are fully informed, and have expressed their consent freely, specifically and unambiguously.

The situation may differ depending on Member States. In some Member States ISPs can, under specific conditions, engage in traffic management policies, for example, to block applications such as VoIP (as part of a cheaper Internet subscription), provided that individuals have given their free, specific and unambiguous, informed consent. Other Member States have chosen to strengthen the principle of net neutrality. For instance, in July 2011 the Dutch Parliament passed a law generally prohibiting providers from hindering or slowing down applications or services on the internet (such as VoIP), unless necessary to minimise the effects of congestion, for integrity or security reasons, to fight spam or in accordance with a court order.[71]

6.5. Conclusion

The analysis of the pros and cons of DPI leads to the conclusion that, on the one hand, its use is inevitable since DPI capabilities are vital for many crucial network services, but, on the other hand, it needs to be used within legal constraints. Rigorous legal protection is required to prevent the potential misuse of such a powerful technology (Norton, 2013); without strict legislation, this sort of deep data filtering can significantly impair user's ability to stay anonymous online (Warwo, 2012). Legally unregulated DPI may result in "an invisible surveillance creep" (Fuchs, 2013a), in which the limited use of DPI for one purpose (e.g. network management or spam filtering) may creep to other, more privacy-sensitive and controversial purposes, such as targeted advertising, or content monitoring for political purposes. As indicated above, DPI is not a malicious technology by design; rather it should

[71] The reasons reported by the press for such a policy option did not refer to data protection and privacy considerations but rather to reasons related to ensuring that users are not deprived of or are offered limited access to information. So it seems that issues relating to access to information motivated this amendment.

be regarded as a neutral technology whose negative uses are as many as the positive ones. It is not the technology itself which is good or bad but the application DPI is used for (Mochalski & Schulze, 2009). As Mueller (2010) suggests, the focus should be placed on the institutions and their forms, rather than the digital code, in order to understand Internet governance issues, such as network security, as well as the technologies used. Whereas code analysis limits our insight into the politics of Internet governance, focusing on the structures of Internet governance helps understand who is advocating for DPI technology. What should be borne in mind is that security technologies do not only have impacts on privacy and individuals, but they also have impact on the society at large (Office of the Privacy Commissioner of Canada, 2009). Given that various actors' interests are involved in the implementation of DPI, its potential depends on the way technological factors interact with economic, political, legal and regulatory factors. As a general point, the use of DPI in Europe by governments or private entities would have to comply with the following requirements. If personal data is to be processed, it must be done transparently (i.e. by informing the data subject and seeking her consent before her data is processed), for a legitimate purpose and conform to the principle of proportionality. In addition to the intrusiveness of the monitoring applied, other aspects are also relevant, such as the level of disturbance to the smooth flow of traffic that would otherwise occur. Regulatory guidance is necessary to determine those inspection practices, which ensure the smooth flow of traffic or can be carried out for security purposes, and therefore may not require users' content, such as, for example, the fight against spam. What is more, guidance is needed to explain the permissible technical parameters to ensure that the inspection technique does not entail processing of data disproportionate to its intended purposes (European Data Protection Supervisor, 2011). The question to be assessed from a data protection and privacy perspective is whether a wait-and-see policy is sufficient. While the data protection and privacy framework does, at the present time, foresee some safeguards especially through the principle of confidentiality of communications, it appears necessary to closely monitor the level of compliance and issue guidance on several aspects that are not particularly clear. In addition, some thoughts should be put forward as to how the framework could be clarified and further improved, in the light of technological developments. If the monitoring reveals that the market is evolving towards massive, real-time inspection of communications and issues related to complying with the framework, legislative measures will be necessary.

CHAPTER 7: Need for an Alternative Regulatory Approach to Cyber Security

2.6. Introduction

The key challenge around which all the current cyber security strategies revolve is related to the dual objective – social and economic - of cyber security policy making. Cyber security policies around the world share the common goals of social and economic prosperity aiming at protecting the society against cyber threats while preserving the openness of the Internet as a platform for innovation and further development of the digital economy (OECD, 2012). These goals are best reflected in the U.S. Cyber Policy Review, which refers to "the dual challenge of maintaining an environment that promotes efficiency, innovation, economic prosperity and free trade while also promoting safety, security, civil liberties and privacy rights" (U.S. White House, 2009). Similar is the approach adopted by Germany, which stresses that ensuring cyber security involves promoting both economic and social prosperity (German Federal Ministry of the Interior, 2011). Special emphasis is placed on the economic aspects of cyber security; therefore, the primary aim of most cyber security policies lies in fostering online trust in order to create the conditions for the Internet to drive economic prosperity. Lack of trust in cyberspace is often cited as one of the most significant obstacles to the use of the Internet and e-commerce (OECD, 2012). According to the UK Cyber Security Strategy, the potential reduction in confidence in online communications could cause "serious economic and social harm to the UK" (UK Cabinet Office, 2011). Both the UK and Spanish strategies, for instance, point out that the development of a safe cyberspace could give a competitive advantage to each country and enable them to maximise the benefits of their digital economies (Spanish Government, 2011; UK Cabinet Office, 2011). As mentioned in the Australian strategy, the future prosperity of the country is closely intertwined with the trust businesses and consumers have in its digital economy (Australian Government, 2009).

Given that the Internet remains fertile ground for an expanding range of beneficial commercial activity, the importance of technological innovation for the digital economy is incontestable. Technological innovation constitutes the primary driver of efficiency in such a competitive environment as it leads companies to develop services or products that better

meet consumer needs. As a means of promoting innovation in technology and security, most countries recognise the need for flexible policies tailored to the volatile nature of cyber security technologies. The Canadian strategy stresses the need for dynamic security concepts able to allow improvements that can best meet emerging threats (Government of Canada, 2010), while the UK strategy also promotes a "flexible cyber security response" (UK Cabinet Office, 2009). The requirement for policies easily adaptable to technological innovation is highlighted by the Japanese strategy, which advocates the implementation of active rather than passive security measures (Japanese Information Security Policy Council, 2010). Promoting innovation entails making the larger business climate amenable to change and, as a consequence, cyber security standards imposed by governments are often viewed as onerous burdens that can only hamper the process of innovation. As a matter of fact, the Netherlands has chosen to address the changing environment of Internet technologies by encouraging self-regulation wherever possible, considering legislation only as an alternative option when the former does not work (Dutch Ministry of Security and Justice, 2011). The rapid pace of change in technological developments often challenges the regulation, which by its nature lags behind market developments.

This chapter examines the feasibility of the development of a regulatory system able to align the private sector's interests with those of individuals while abstaining from impeding technological innovation. The shortcomings of a technology specific approach to cyber security, coupled with the strong objections on the part of governments to introduce regulatory-based security standards, call for the need to adopt a regulatory approach imposing security obligations that are broad enough in terms of technical requirements to accommodate all types of technological designs, and also clear enough to provide essential guidance to data controllers, thus safeguarding data protection. To this end, the role of performance standards is analysed and suggestions are made as to the integration of industry's practices, such as the adaptive management process, into the law-making process as a potential means of striking the right balance between two seemingly irreconcilable concepts – regulation and innovation. The concept of smart regulation is also examined and emphasis is placed on the need for its integration through the entire law-making cycle.

2.7. Need for a Regulatory System That Promotes Technological Innovation

7.2.1. Objections to regulatory-based cyber security standards

There has been considerable objection to regulation establishing cyber security standards on the grounds that government mandates would stifle innovation in this field. The most frequently articulated arguments against regulatory-based private sector cyber security standards are the following. First, imposing specific forms of technologies on companies would deprive them of economic incentives to search for better approaches to enhancing cyber security, since it would prevent the uptake of new technologies that could provide better outcomes (IPCC, 2007). The development of more effective security technologies is primarily motivated by economic incentives, which are the key drivers of technological innovation (Financial System Inquiry, 2014). Second, regulation mandating certain technologies would actually hamper cyber security and impede companies' flexibility due to forcing them to introduce cumbersome or inefficient security measures (Dilanian, 2012; Etzioni, 2014). The third argument is related to the fear of compliance, which is likely to block technological experimentation (Lewis, 2009). The fact that companies might be discouraged from changing or improving systems and processes could lead them to adopt a conservative view in order to minimise regulatory compliance risks or out of fear that regulatory standards will be tightened again (Financial System Inquiry, 2014). Moreover, overly prescriptive regulatory approaches that increase technical barriers to trade risk freezing security solutions and constraining innovation in security by hindering the ability of companies to implement globally consistent security techniques. The adoption of specific regulatory technical requirements would result in "balkanising the Internet into different markets with different technical regulation" creating a fragmented legal framework with unpredictable rules which would frustrate innovation and harm economic competitiveness globally.[72]

Another concern raised by the opponents of technology specific regulation is the cost of an overly restrictive regulatory plan, which could be a reduction in security. Cyber security

[72] A concern expressed by the Internet technical community - *See* Non-Governmental Perspectives on a New Generation of National Cybersecurity Strategies: Contributions from BIAC, CSISAC and ITAC, in OECD, 2012, Cybersecurity Policy Making at a Turning Point: Analysing a New Generation of National Cybersecurity Strategies for the Internet Economy, p.86, available at: http://www.oecd.org/sti/ieconomy/cybersecurity%20policy%20making.pdf.

regulations would impose substantial costs, which the private sector would be incapable of meeting profitably. As regulatory requirements are viewed as "unfunded mandates" unfairly and inappropriately adding security measures beyond those introduced by private organisations, corporations argue that they should be compensated by the government for resulting costs (Etzioni, 2014; Gattuso, 2012). Furthermore, the inherently dynamic nature of technology often allows regulatory schemes even less of a chance of keeping up with technological advances. Both the emerging threats and technologies move at a faster pace than regulators' ability to establish and maintain effective rules; technology changes so rapidly that when a legislation adapts, it is time to adapt again (RSA, 2014). Irrespective of how carefully a regulation is designed, regulators cannot have the knowledge to craft solutions as it is difficult to foresee all the new products and delivery mechanisms that technology might enable. But wrong regulations could be counterproductive to the efforts of ensuring cyber security since they can hurt security by blunting private-sector innovation and flexibility (Gattuso, 2012). Finally, rather than reflecting technological needs, regulations are often biased, subject to political pressures, while the lengthy regulatory processes can also create uncertainty for the industry during the transition period increasing the risk of unintended consequences.

The idea prevailed among the private sector, that regulation should avoid unjustifiably inhibiting innovation by creating trade barriers, has as a result corporate leaders to adhere to the laissez-faire and libertarian principle that the private sector has a right to be let alone by the government to independently determine what kind and level of cyber security it needs (Etzioni, A. 2014). It is argued that cyber security policymaking challenges can only be successfully addressed if innovation in technology and security are allowed to advance.[73] Therefore, a 'light-touch approach to regulation' has been proposed, according to which companies should be encouraged to follow voluntary codes containing security best practices rather than top down prescriptive measures that risk unnecessarily interfering with the development and deployment of security systems (U.S. Department of Commerce, 2011). Certain industries, which are important to innovation and economic growth, are more likely to be responsive to flexible structures in order to promote security that is in their own interest by solving the unique security problems they face. In this regard, the role of government should be limited to developing protections that advance innovation and enhance cyber security. This

[73] *See* Response from Business and Industry Advisory Committee (BIAC), OECD (2012) *Ibid.*70.

idea is strongly supported by a variety of think tanks who warn against European data privacy restrictions that could stifle innovation and restrain economic growth in the region (Westervelt, 2014). Likewise, it has been argued that by mandating cyber security measures, "the U.S. may end up hobbling its strongest weapons in the war against cyber threats" (Gattuso, 2012).

7.2.2. Failed attempts to mandate cyber security standards

The strong private sector opposition to regulatory cyber security standards has led governments to abstain from mandatory regulation that would be considered burdensome by private organisations. The case of the U.S. Congress is illustrative of the government's tendency to protect the interests of private actors by rejecting the introduction of regulations that encounter resistance by companies and is reflected in the following statement of President Obama: "My administration will not dictate security standards for private companies".[74] The first cyber security proposals were made by Richard Clarke, a former National Coordinator for Security, Infrastructure Protection, and Counter-terrorism for the United States, who attempted to implement an ambitious regulatory regime during his tenure, but with no success as it was blocked by anti-regulation forces. Recognising the limits of what could be done in practice, Clarke called for the introduction of an elementary cyber security measure that was low-cost and high-yield and could have helped secure the vast majority of information transmitted on the Internet – setting filters on the major ISPs, where nearly all Internet traffic passes through, in order to detect malware and cyber attacks with no noticeable delay in the connection speed. However, business interests also prevented the implementation of this proposal. Similar has been the fate of any piece of legislation that has attempted to either introduce mandatory minimum standards for the private sector or to sanction private sector entities that fail to adopt reasonable data security practices (Clarke & Knake, 2010; Etzioni, 2013).

[74]*See* Dilanian, K., 2012, U.S. Chamber of Commerce leads defeat of cyber-security bill,*Los Angeles Times,* August 3, available at:http://articles.latimes.com/2012/aug/03/nation/la-na-cyber-security-20120803.

7.2.3. Need for a widely accepted obligation

Given that the private sector has strong incentives to invest highly in cybersecurity due to competition, one might well expect that corporations – as rational actors - should be keen to voluntarily implement the necessary security measures. After all, they suffer the most from cyber attacks, which harm their equipment and facilities and cause considerable damage including economic losses, losses of trade secrets and damage to their reputation. Therefore, it may seem obvious that companies would do their best to protect their trade secrets and hence their profits. However, the security level adopted by the private sector in cyberspace remains far from satisfactory, as evidenced by the fact that companies are increasingly inundated by security breaches. Even though millions are invested in cyber security, companies continue to bring products to the market to which hardly any thought seems to be given towards cyber security providing new opportunities to cyber criminals, and thus perpetuating the problem of cyber insecurity.

The reasons for private sector's reluctance to take the necessary measures are varied but are primarily economic. Most CEOs seem to attach greater importance to short-term costs and benefits to the detriment of longer-term effects, whose consequences often take years to unfold, while others tend to underestimate even the short-term consequences of weak security policies (Kahneman, 2011; Yadron, 2014). This problem is often compounded by executives' lack of technical knowledge, which is reflected in the following analogy between cyber security and environmental law made by a cyber security expert: "Cyber-security resembles environmental law in that both fields are primarily concerned with negative externalities. Just as firms tend to underinvest in pollution controls because some costs of their emissions are borne by those who are downwind, they also tend to underinvest in cyber-defences because some costs of intrusions are externalised onto others" (Sales, 2013).

As demonstrated above, the faith-based assumption that the private sector will act in good faith, and hence will take all the necessary measures in order to secure its networks and computers, has proved inadequate to provide the required level of security in cyberspace. Many years after the emergence of the Internet, this faith-based approach has indicated that merely relying on the companies' willing adoption of strong security techniques will by no means address the problem of cyber-attacks (Lewis, 2009). As stated by a long-time proponent of deregulation, Christopher Cox, a former chairman of the U.S. Securities and

Exchange Commission, “the last few years have made it abundantly clear that voluntary regulation does not work.” The failure of private actors to recognise their responsibilities illustrates what is missing at present – a jointly accepted obligation to ensure the creation and maintenance of a more secure cyberspace (Veraart, 2013). This goal, however, cannot be achieved unless standardised best security practices are agreed upon and continuously implemented by diverse security companies (OECD, 2012). Taking into account the lack of consensus as to what is considered to be a best security practice, the limitations of self-regulation call for a different approach to cyber security.

Another concern deriving from self-regulation is associated with the fact that decisions regarding the level of cyber security, which often refers to the protection of (sometimes sensitive) personal data, are increasingly made by private organisations with almost no public awareness or accountability. The question then becomes how to ensure optimal and consistent participation when cyber security is entrusted to a large extent to private actors (Dupont, 2013). It is noteworthy that in the early 2000s, the aim of cyber security was to create the conditions for the Internet economy to become a significant source of economic growth by fostering trust online; as expected, the lead role was then given to the private actors with the required technical expertise. Even though the aforementioned objective has nowadays been achieved, and in spite of the increasing number of cyber threats, the Internet continues - after so many years - to be “driven more by considerations of interoperability and efficiency than security” (U.S. White House, 2009).

The main danger posed by regulating through code is the fact that issues with enormous impact on all members of society are addressed by a privileged few, whose decisions are often made outside the public eye. In contrast with the regulation through law, which follows a democratic procedure, or through social norms, which takes effect only when a significant portion of the public is aware of the changed norms, the technological regulation is far more discreet. Developing technologies are kept out of the public eye, mostly due to trade secrets and intellectual property law, until the point they are released on the market and start affecting the prevalent regulatory scheme. Moreover, the principal goal of a private organisation is to make profits from the development of technologies focusing primarily on the economic benefits of cyber security and therefore its interests are rarely in line with those of society at large. Granting such a great power to private parties without making them answerable to any

democratic process results in sacrificing some of the most important rights, such as the right to privacy and security (Kumar, 1998; Lessig, 1999).

As indicated above, addressing the problem of cyber insecurity requires a policy evolution from the mind-set that drove the growth of the Internet and focused mainly on the individual interests of specific participants; what is required is a different approach to cyber security that aim at the protection of the society as a whole. Therefore, a body answerable to the public for its actions is required to maintain some role in either making or overseeing technical decisions that shape regulation affecting the entire society. Governments can play a lead role in the implementation of best practices by introducing policies and legislative requirements that would provide a clear direction to economic and social actors and elicit actions that the private sector would not otherwise perform (OECD, 2012). Although the lack of regulation was viewed as beneficial in the past, government intervention is now considered crucial in response to market failure; regulation has its own limitations but it might be preferable to inaction which has proved inadequate to address cyber insecurity.

Commenting on the U.S. President Obama's cyber security executive order[75] that authorised the creation of an incentivised set of voluntary security guidelines for the protection of networks connected to critical infrastructure facilities, James Lewis, a highly regarded cybersecurity expert at the Centre for Strategic and International Studies pointed out that if the private actor does not act on the guidelines, the rules should become mandatory rather than being solely voluntary as there is always a limit to what can be done by organisations that are not covered by regulation (Lewis, 2009). Related is the concept of 'responsive regulation', which has been suggested as an alternative to the dichotomy of state regulation versus self-regulation (Ayres & Braithwaite, 1992). The proposed default strategy consists in allowing discretion to private actors to find the optimal ways to achieve regulatory goals based on their particular needs and capacities, but in case those actors fail to implement effective strategies, the government retains the ability to escalate its level of interventionism by shifting to command-and-control regulations that involve various forms of punishment. Responsive regulation principles are considered to be inherently compatible with the need to foster

[75]*See* Westervelt, R., 2013, President Authorizes Cybersecurity Plan for Critical Infrastructure, *CRN*, February 13, available at: http://www.crn.com/news/security/240148452/president-authorizes-cybersecurity-plan-for-critical-infrastructure.htm.

innovation in a highly competitive business context as it advocates delegated and non-intrusive regulation that is more likely to generate cooperation and innovation.

7.2.4. Regulation & innovation: two irreconcilable concepts?

As aforementioned, the opponents of regulation in cyber security, who do not similarly oppose regulation in other fields such as copyright or intellectual property law, object to government mandated cyber security standards on the grounds that regulation is ineffective, imposes unjustifiable financial burdens on the companies and hinders innovation. First, it is argued that regulation is ineffective for improving cyber security because many regulated industries have not proved to be secure. But it should be borne in mind that irrespective of how well designed a regulation might be, there are also other parameters, such as weaknesses in law enforcement or the general insecurity of the cyber environment, that would undermine the effectiveness. In this aspect, it is noteworthy that cyber security has for a long time been a secondary concern for regulatory agencies, who might lack the interest and expertise to improve cyber security, and thus they consider it as peripheral to their mission (Lewis, 2009). The main question here should be whether regulated industries are more secure than they would have been otherwise; for instance, banks still suffer losses but their losses would be much more severe if they had not been forced by regulators to enhance security. Furthermore, in terms of the costs imposed by regulatory-based standards, they are viewed as disproportionately high by private companies, which claim that the implementation costs would be lower if a firm was given some discretion in how it meets its security target. However, it becomes clear that costs should not be an obstacle for large companies if one compares the amount of money invested in cyber security with the size of many companies' revenues.[76]

As stated by Lewis (2009): "The intellectual heritage of deregulation lives in assertions such as any regulation to improve security will hurt innovation. Like all lobbyist mantras, it contains a grain of truth while being fundamentally and dangerously wrong. Innovation is a complex process, and simple statements about cause and effect deserve only scepticism".

[76] For example, the revenues of a company named Target, which suffered a major breach in 2013, topped $72 billion that year, while its investment in cyber security was $1.6 million, which equals roughly 0.0002% of its revenue.

Taking into account the diversity of factors involved in the process of innovation, it is difficult to single out regulation as the main reason for hindering innovation. Besides, the absence of regulation in cyber security for a long period of time seems to show exactly the opposite; despite the lack of regulatory standards, the wave of innovation expected since the emergence of the Internet has not actually been produced (Dupont, 2013). A large number of examples, where governments imposed safety regulations and mandated action by the private sector to secure the public, indicate that governments can instead facilitate innovation through implementing regulation. The regulatory requirements to manufacture, for instance, safer aeroplanes or cars have not stifled aircraft or automobile innovation. On the contrary, they have promoted competition by creating demands for new and safer products while encouraging companies to work in a conceptual framework that makes security a primary concern. The case of the automobile industry is illustrative of the extent to which regulation can both inspire innovation and foster security, as safety requirements have now become a useful marketing tool, whereas in the past automobile companies resisted innovations such as seat belts or safety glass, thanks to which traffic fatalities have nowadays significantly declined. Another example associated with the use of the Internet refers to the safety regulations with respect to online transactions; by protecting consumers using electronic payment methods, these regulations have played an important role in developing the electronic payment market since they have boosted consumers' confidence in electronic payments leading to their mainstream adoption (Lewis, 2009).

Whereas an overly prescriptive regulation that mandates certain technologies or forbids certain activities can be an obstacle to innovation, careful regulatory drafting can actually avoid risk to innovation. At this point, it is worth briefly explaining the categories of regulatory standards, that is, standards that specify with a certain degree of precision the actions that need to be undertaken for specific objectives to be achieved, and consist of two broad classes – technology and performance standards. Technology-based approach to cyber security involves the regulator stipulating the security techniques and products that should be used by specifying the technological steps a firm should take to secure its networks and computers. On the contrary, more general standards are introduced by the performance-based approach, which expands compliance options beyond a single mandated technology and focuses on the outcomes to be achieved (IPCC, 2007).

A regulatory agenda for cyber security does not need to prescribe actions in exhaustive details. Instead, a well-designed regulation should establish common performance baselines and metrics for an acceptable level of security. Such a regulation could prove to be particularly helpful for the private sector as not only can it provide a reasonable approach for companies that "lag behind in establishing security measures" but it can also address specific elements of security where private actors have no reach or control (Westervelt, 2014). By laying out the regulatory goals and principles, assigning principles and creating processes to ensure compliance, regulation can both promote cyber security and avoid hindering innovation. In this regard, the challenge faced by the regulators is how to balance the benefits against the risks of regulatory flexibility without undermining the required level of protection. In pursuing a performance-based approach, governments need to consider the text of regulation so that it does not entail ambiguity and interpretation difficulties. Even though a regulation that only describes the outcomes is flexible to adapt to the future and does not create obstacles to the innovative process, it may result in legal proposals lacking the necessary safeguards that lead to overbroad and vague powers posing an unjustifiable threat to privacy and security.

Managing the complexity of pursuing the double objective of cyber security policy making – economic and social prosperity – is currently viewed as one of the main challenges for policy makers. The complexity lies in the fact that these two objectives are seemingly contradictory. On the one hand, technological innovation, which is a powerful force for competition and the key driver for the development of more effective products, constitutes a significant prerequisite for the economic growth of any country. On the other hand, social prosperity can primarily be attained through regulations that protect individuals by providing the necessary safeguards. In terms of cyber security safeguards that could protect the members of society against the interests of the privileged private actors, who often fail to implement robust security techniques thus jeopardising individuals' personal data, there has been strong objection to government mandated standards on the grounds that they would stifle innovation. This laissez-faire principle that companies should be let alone to independently determine what security measures are more suitable to their business model has resulted in governments abstaining from introducing regulatory-based standards and instead resorting to cajoling the private sector to adopt voluntary security measures.

The failure of the free market approach, as evidenced by the increasing number of security breaches, compounded by the severe risks posed to individuals' personal data by weak security policies, dictates the need for standardised best security practices to be continuously implemented through regulation. Regulation may not solve the problem of cyber-attacks, but it definitely is a more sustainable option compared to the expectation that the laissez-faire solutions will miraculously produce enhanced cyber security. Analysing whether regulatory standards or market-based instruments are preferable, it has been suggested that the most appropriate approach for developing countries is a transitional approach, whereby technology standards are first introduced, followed by performance standards, and finally by experimentation with market-based instruments (Russell &Vaughan, 2003).

In contrast to the idea prevailed among the private sector, regulation and innovation are not inherently at odds; by contrast, regulation can inspire innovation to supply a public good by creating markets for safety. A symbiosis between governments and innovators can be achieved provided that governments strike a balance between the requirements of public safety and growth; a balance that allows the benefits of innovation to blossom while maintaining stability. As stated by Lewis (2009), "Too much regulation will kill economic growth, too little will put the country at risk". The lack of government intervention would damage the public interest as much as over-regulation. What is needed is a change in the political mindset which will allow governments to set cyber security standards to protect individuals' rights to security and privacy, rather than being responsive to corporate lobbying. The most crucial step, however, is to make cyber security a primary mission for regulatory agencies, who need to understand the centrality of cyber security and thus ensure the delivery of reliable and safe services to the public.

7.3. Need for a Regulatory System That Aligns Companies' Interests with Individual Users' Interests

7.3.1. Performance-based regulation as a new approach to cyber security regulation

The EU legislation currently does not ensure that the interests of companies and individuals are well aligned when it comes to the security measures implemented by the companies while

collecting, processing or storing personally identifiable information. Most of the times individual users are even largely clueless about the amount of personal data collected and how it is further used due to the obfuscating data processing practices. It is possible that what is good for a private organisation is good for the individual users but there should also exist a regulation ensuring this alignment of interests. The fact that the private sector does not attach the necessary importance to protecting personal data, coupled with the large number of security breaches resulting from the lack of strong security measures, indicate that the law's goal of maximising individual welfare is probably not being met. Moreover, private organisations often reformulate their practices to evade regulation with a speed that will only increase in the device-mediated world of Big data, whereas regulators, limited institutionally to slow and cumbersome responses, can rarely stay caught up for long. Therefore, an alternative regulatory instrument is required, which would intervene in this dysfunctional regulatory cycle and have the potential to make the law as agile as companies are.

The process of establishing performance standards can provide a new lens through which to place individuals at the centre of the law, as it seeks to engineer an alignment of incentives between companies and users when the market fails to do so. The performance-based regulatory paradigm introduces a more adaptive approach than prescriptive regulation since it sets performance goals rather than specifying behaviour. The effective implementation of this type of regulation depends on periodic surveillance and accountability in order for regulators to be able to test the level of performance of the regulated entities and thus suggest amendments to regulation when necessary. Performance-based regulation aims to accomplish two objectives. First, it aspires to unite the interests of companies with the ultimate goal of the regulator. Second, the ongoing monitoring can provide feedback to both marketers and regulators that can be used to improve the marketplace and regulation in a virtuous cycle.

Incorporating performance standards into regulatory goals is by no means a new idea, but in recent years there has been renewed interest in expanding the use of performance standards across different regulatory settings (e.g. health, safety and environmental regulation). In areas such as the environmental regulation, prescriptive regulation has been supplemented with performance-based regulation monitored through ongoing field-testing. The role of the law is to set limits on a company's emission while it allows the company to determine how to meet

those limits.[77] Hence, regulators are able to both harness the company's innovation in finding ways to meet the standards and also obtain systematic information by continually monitoring the company's performance, which help them prevent substantial harm and propose more effective regulation. In an attempt to apply the performance-based regulatory paradigm to cyber security regulation, the following analysis clarifies the meaning of performance standards, describes the strengths and weaknesses of performance-based regulation, and identifies the likely conditions for the effective use of performance standards.

7.3.2. Defining performance standards

The primary aim of any kind of regulation is to improve the performance of regulated entities in order to reduce social harms and thus benefit the society as a whole in the long run. In this sense, the term 'performance-based regulation' may seem redundant since regulation should by definition aim at ameliorating the behaviour of an individual or an organisation. However, regulators can direct the regulated entities to enhance their performance in at least two basic ways – by specifying either the actions to be taken or the desired level of performance to be achieved. Prescriptive regulation is typified by concrete rules that dictate how particular technologies, procedures or processes should be designed so as to push the regulated behaviour toward the regulator's ultimate goal. On the contrary, performance-based regulation sets a general standard that is closer to the regulator's ultimate goal and allows the targets of regulation to decide how to meet that standard. The underlying premise of performance standards is the desire to let the regulated entities reach the appropriate level of performance in the most cost-effective way possible. Adopting a risk-informed approach to achieving the goal set, performance-based regulation specifies performance objectives that better describe the outcome required while the concrete measures that need to be applied are left to the discretion of each regulated entity.

In order to analyse effectively the potential and limitations of performance-based regulation, it is significant to fully comprehend its meaning by clarifying the definition of the term 'performance' and identifying the several distinctions of performance standards. To begin with, in performance-based regulatory systems the notion of performance can be interpreted

[77] *See* Regulation (EC) No 443/2009 of the European Parliament and of the Council of 23 April 2009 setting emission performance standards for new passenger cars as part of the Community's integrated approach to reduce CO2 emissions from light-duty vehicles.

in different ways; for instance, performance can be used as a basis for assessing regulatory agencies, as a criterion for allocating compliance and enforcement resources, or as a trigger for the application of tiered regulatory standards. Although the above uses of the term 'performance' frequently arise in policy and academic discourse, the most common conception in the literature of policy instrument choice is that of performance as the basis for the legal commands found in regulatory standards (Daly & Foushee, 1981; Coglianese et al., 2003; Lovells, 2014). In other words, performance standards introduce regulatory requirements so that the regulated entities demonstrate a satisfying level of performance.

Furthermore, the distinctions of performance standards vary depending on the type of the problem the standard aims to solve; the underlying basis for the threshold reflected in the performance standard; the scope of the regulation's ultimate goal; and the specificity of the regulation. As it will be indicated below, the last criterion is the most relevant in the context of this chapter. With regard to the specificity of regulation, performance-based regulation can specify the standards either in a loose or a tight manner. Loosely specified standards provide the regulated entities with more discretion but with less guidance. Whereas tightly specified regulation employs quantitative measures of performance, loosely specified regulation calls for regulators to make qualitative judgements. The distance between the performance goals set by regulation and the ultimate objective that leads to the creation of a regulation determines the amount of flexibility embodied in performance standards. Based on this criterion, performance standards are set either according to the level of performance that is achievable taking into account the known technologies or according to the appropriate level of risk determined (Coglianese et al., 2003). In this context, a performance standard can, for instance, simply introduce a broad societal objective (such as preventing security breaches) or specify a narrower or subsidiary goal (such as requiring a certain level of security).

As mentioned above, the discretion granted to the regulated entities offers them the required flexibility in seeking the lowest-cost means for achieving the stated level of performance while meeting the performance standard in a way that best suits their needs. Not only does performance-based regulation promote innovation since it abstains from specifying the design requirements of a product or a process, as opposed to prescriptive regulation, but it also promotes sustainability since performance standards can accommodate technological change and the emergence of new hazards in ways that technology-based standards cannot (Hemenway, 1980; Besanko, 1987). The performance-based approach to standard setting aspires to align the goals of the regulated entities with the regulator's ultimate objective. In

doing so, it provides regulated entities with greater ability to adapt to changes in both the environment and available technology, while it enables them to obtain greater knowledge of their own processes and offers them greater facility for experimentation. In the context of private organisations, it is argued that combining performance-based regulation with market mechanisms allows firms with lower compliance costs to trade with those with higher compliance costs and thus results in reducing total costs while promoting innovation (Coglianese & Lazer, 2003).

Yet, performance-based regulation is not without limitations. The most serious shortcoming of this type of regulation lies in the fact that loosely specified performance standards are often imprecise and, as a consequence, create uncertainty for both regulators and regulated entities in terms of compliance and enforcement. This problem appears to be more serious when the regulated entities are small businesses which may not be able to afford the excessive costs involved in searching for ways to meet broadly defined regulatory standards. Issues related to lack of clarity also arise in the case of tightly specified performance standards because setting optimal quantitative thresholds requires a detailed understanding of the relationship between the technologies available and the regulator's objective, which is frequently poorly understood. The fact that performance standards are based on predictive model-technologies that are known to work well – may lead to "legitimate self-delusion" on the part of regulated entities, which may erroneously interpret their models in such a way that makes their approaches seem to perform well (Coglianese et al., 2003). In such cases, when standards are based on predictions rather than actual measurable events, distinct challenges arise in measuring the level of performance (May, 2004). Moreover, as far as the regulated entities' discretion is concerned, it may actually be significantly constrained especially when performance standards are narrowly defined. But even in the case of broadly defined standards, strict adherence to highly specified standards might be required thus resulting in many of the downsides of a prescriptive regulation and calling into question the desirability of performance standards versus design standards (Daly & Foushee, 1981).

7.3.3. Requirements for a successful performance-based regulation

In order for performance standards to fulfil their potential in aligning the actions of the regulated entities with the regulator's ultimate objective, they should be applied under specific conditions. Otherwise, the discretion granted to the regulated entities presents the possibility of engendering risks, since the wide margins of interpretation involved in

performance-based regulation may result in undesirable side effects. First and foremost, given that the law itself lacks essence unless efficient enforcement mechanisms are put in place, particular importance should be attached to enforcing compliance to performance standards. At this point, it is crucial to highlight the need for a change to the prevailing approach to enforcement. Instead of agencies taking action only after it has been proven that the regulated entity failed to abide by the rules, it would be better to intervene before an undesirable event occurs. Enforcing performance standards only ex post may be too late to avert a problem and may result in systematic under-enforcement. Performance-based regulation often requires the application of the so-called "performance indicators" that demonstrate the frequency of problems caused as well as the likelihood of such problems to occur again (Coglianese et al. 2003). To provide agencies with enough time to prevent bad performance, performance indicators ought to be embedded well below the level of the regulator's ultimate objective. For instance, an agency can develop performance indicators using probabilistic risk assessment, which means that the agency can assign certain risk levels to each company based on periodic reviews of the company's level of performance, while it takes progressively more control over facilities with higher risk levels; if a company, for example, receives two consecutive risk ratings, it is asked to propose corrections.

Another significant condition for a successful performance-based regulation is the ongoing interaction among key stakeholders. The dialogue between government and stakeholders, both at the time of setting the standards and at the time of application, can be used to inform and educate the latter with respect to the appropriate implementation of performance standards, thereby making them aware of what actions they should take or avoid in order to meet the standards. When government engages with industry, for example, it becomes apparent how difficult it is for small companies to act in accordance with performance-based rules due to the lack of clarity involved in them. In addition, dialogue helps expand the set of possibilities available to regulators by creating industry-wide criteria for assessing performance. Dialogue can also have beneficial results for consumers since it can raise consumer awareness allowing to assess whether individual terms or products are unfair, deceptive or abusive and thus make more informative decisions (Coglianese et al. 2003).

Furthermore, the third condition to be met reveals the reason why governments avoid relying extensively on performance targets. Even though performance-based regulation presents the advantage of decentralising government's responsibilities by providing greater flexibility to the private sector, the appropriate implementation of performance standards depends on the

ability of government agencies to specify, measure and monitor the performance of the regulated entities, which is not always an easy task given that reliable information about performance may sometimes be difficult or even impossible to obtain (May, 2004). It is crucial that government gets so actively involved in this process that it eventually seems that it is essentially running almost everything; otherwise performance-based regulation will function poorly under the wrong conditions.

Effectively assessing compliance with performance standards requires ongoing periodic testing of the regulated entities' levels of performance. In terms of the timing of compliance testing, however, performance-based regulatory systems vary in operation, since the testing of compliance can be carried out before the performance, which is the regulation's real target, after the performance or at regular intervals during the performances. In order for performance-based regulation to function properly, compliance should be measured regularly after performance standards have been put in place (Willis, 2015). The lack of ongoing surveillance and accountability can create severe problems in a performance-based regulatory system. When performance is assessed immediately after the introduction of performance standards, it cannot capture behavioural patterns and therefore the regulator's goal can merely be met in the short term but not in the long run. Such is the case, for instance, of modern building codes, which measure compliance with performance standards only at the time of installation, whereas there is no further obligation for ensuring performance over time (e.g. testing whether the materials continue to perform as desired or whether substantial problems have been caused by leaky buildings) (May, 2004).

Periodic performance testing proves to be of vital importance for performance-based regulation to fulfil its potential, as it does not only reveal which regulated entities are not meeting the regulatory goals but it also reveals which regulations are failing to meet the regulatory goals. Policymakers and citizens currently lack sufficient information regarding how the marketplace is actually functioning as well as the effectiveness of consumer law. Despite the fact that it is nowadays a common practice for companies to collect information on how their customers use their products, regulators obtain this information solely on a sporadic and partial basis (through surveys, reports, or consumer complaints) and, as a consequence, the information they have is inadequate to formulate policy based on what is really happening in the marketplace. Yet, to develop policies that are grounded on data corresponding to the reality and not on regulator's theoretical models, regular field-testing is necessary to provide regulators with data that would level the regulatory playing field and

increase regulator accountability while empowering the role of individuals both as consumers and citizens (Willis, 2015). As aforementioned, periodic performance testing would also allow citizens to decide whether a regulation has succeeded at achieving its objective or not. As long as the results of performance testing were made publicly available, they would raise awareness and understanding of the strengths and weaknesses of a regulation, and would potentially change citizens' behaviour motivating them to actively engage in the political process by putting pressure on regulators to improve regulation.

7.3.4. Lessons from performance-based Consumer Law

In an attempt to bring consumer transactions in line with consumer expectations and make the law as nimble as private organisations are, Willis (2015) suggests a new approach to consumer law which introduces performance standards designed to evaluate actual consumer comprehension (comprehension standards) and consumer product choices (suitability standards). Comprehension and suitability standards are the two ways in which the performance-based regulatory paradigm can be applied to consumer transactions. By setting performance standards relating to companies' activities, performance-based consumer law aims to incentivise companies to educate consumers instead of obfuscating their practices and motivate them to develop products that are suitable for the consumers' circumstances. The policy structure proposed is of particular interest since it can be applied in several areas, such as the field of personal data protection. To start with, the role of comprehension standards is to empower consumers to actively engage in the design of marketplace products by maintaining their autonomy in determining which transactions they would benefit from. The rationale behind the concept of comprehension standards is that increasing consumer awareness with respect to the individual product features and the differences among products would radically change companies' incentives. Comprehension standards can capture the effects of companies' marketing practices and therefore private organisations would be incentivised to abstain from misleading practices whose goal is to outsmart the consumer; instead, companies would be motivated to educate consumers, whereas savvy consumers may even push companies to design products that best meet their needs.

In contrast to regulators, who inform consumers solely through public-education campaigns, the private sector is best situated to perform this task at a lower cost and through changing conditions. Hence, the role of performance-based regulation should be to encourage private organisations to implement welfare-enhancing practices through the introduction of

comprehension standards. Given that companies collect and analyse their customers' information for marketing and product development purposes, they have expertise in educating consumers. What is more, companies are even able to tailor marketing to individual consumers' real-time circumstances thanks to the advanced technologies available in the device-mediated age of Big data. Large companies already run thousands of randomised studies annually[78]; for instance, Facebook alone runs over a thousand experiments every day.[79] Based on these studies, marketers are in a position to respond much faster than governments to evolving consumer beliefs and also affect consumer behaviour. What is currently missing though is what performance-based regulation aspires to achieve; to shift companies' focal point to consumers' needs and make them use this knowledge and tools not only for their own profit but also towards increasing consumer awareness.

In addition to comprehension standards, consumer transactions can also be regulated by another type of performance standards – the suitability standards – which are particularly useful in cases where the cost of achieving comprehension exceeds its benefits. Suitability requirements define which uses of a product would be suitable or not and aims at increasing consumers welfare by either requiring companies to develop products embedding certain attributes beneficial for consumers or by eliminating product features that result in more costs than benefits for consumers. Instead of testing the suitability of each transaction, which would be excessively costly, or enforcing ad hoc compliance, which would discourage companies from taking any action, performance-based regulation sets performance benchmarks regarding the proportion of a company's customers that must use its products suitably and then field-tests a sample of the customers to assess whether the standards are met.

Suitability standards help mitigate the two main problems arising from prescriptive regulation; over-inclusiveness and under-inclusiveness. First, suitability standards can motivate companies to develop methods to channel the right products to the right consumers, while by providing flexibility they avoid the innovation-retarding effects that prescriptive regulation is likely to have. Second, the higher level of generality in suitability standards renders the law-making process more agile compared to prescriptive regulation and enables

[78]*See* Obama, B., 2014 Economic Report of the President, March, available at https://www.whitehouse.gov/administration/eop/cea/economic-report-of-the-President/2014.
[79]*See* Bakshy, E., Eckles, D. and Bernstein, M., 2014, Big Experiments: Big Data's Friend for Making Decisions, April 3, available at https://www.facebook.com/notes/facebook-data-science/big-experiments-big-datas-friend-for-making-decisions/10152160441298859/.

regulators to respond more quickly to changes in product offerings that make the use of a product unsuitable. Not only can suitability standards mitigate the weaknesses of prescriptive regulation, but also have the potential to suggest improvements in design standards. Establishing suitability standards can lead to a consensus on appropriate uses of a product that can in turn result in more focused product-design regulation. The flexibility provided by performance-based regulation allows companies to experiment more easily and thus discover changes in the design of a product that regulators will look at examining what product designs are feasible and based on this knowledge they will then be able to introduce industry-wide design standards.

As it has been indicated, successful performance-based consumer law can provide companies with the incentives that are currently missing and direct them towards welfare-enhancing practices. The relationship between the regulator and the marketers is reciprocal; on the one hand, performance standards offer the necessary discretion to companies to decide the best way to meet the benchmarks set and, on the other hand, companies' choices provide regulators with market-tested product designs on which to base design-regulation decisions. Performance-based approaches can be integrated into incentive-based regulation either directly (positive incentives) or indirectly (negative incentives). As far as negative incentives are concerned, to encourage continuous improvement of the companies' practices, regulators could, for example, charge a fee for behaviours that increase risk or impose liability for unfair, deceptive or abusive conduct. If companies had greater responsibility for defective or insecure products, their incentives would be to design products that could be used only in welfare-enhancing ways; this form of liability can prove crucial in an era when technology facilitates the sale of products with unsuitable features or without any consideration for security features.

7.3.5. Incorporating performance standards into cyber security regulation

The current cyber security landscape calls for a regulatory instrument capable of providing data controllers with the necessary incentives for adopting security practices having as its primary goal the protection of individuals' personal data. As evidenced by the increasing number of security breaches, the technology neutral approach adopted by the relevant EU legislation seems to have failed to achieve this goal, that is, motivating companies collecting and storing personal data to take data protection more seriously. In an attempt to avoid the innovation-retarding effects of a technology-based regulation and ensure the sustainability of

the law, EU regulators have abstained from imposing specific security obligations on data controllers providing them with great discretion in interpreting the letter of the law; although the EU laws are always accompanied by clarifying guidelines, the latter have no binding character. Incorporating performance standards into cyber security regulation might result in redirecting the creative potential of the private sector towards the development of welfare-enhancing practices by uniting companies' interests with the regulator's goal, while leaving ample space for technological innovation. Even though performance standards confer on companies both flexibility and responsibility to determine how to meet the legal requirements, they do not give them complete discretion, as is the case of other forms of technology neutral regulation; instead, the standards set an intermediate target that companies must meet. Performance-based regulation aspires to achieve what is currently missing from cyber security regulation; to place users at the centre of cyber security landscape by shifting regulator's focus from the data controller's actions to the effects of those actions on users. In this way, the performance-based approach to cyber security regulation brings a new perspective, which situates individuals as principals, and regulators and data controllers as their agents. This new perspective, however, also requires a new vision of how law ought to be made.

A significant benefit of the performance-based regulatory paradigm lies in the fact that it introduces the notion of "smart governance", according to which regulators have the responsibility and authority to routinely review the performance of regulation and to make constant adjustments to direct regulation closer to its goals (Dale et al., 2007). The process of periodic testing of compliance with performance standards would give the power to regulators to systematically review performance data and would enable them to respond more quickly to them suggesting regulatory improvements. For instance, based on the test results regulators would be in a position to iteratively revisit the question of how insecure products or improperly applied security techniques should be regulated. Hence, the process of law-making would not rely on regulatory and judicial models or myths about the 'average' user or the 'appropriate' security techniques; instead, empirically informed law-making examines what is actually happening and can propose more grounded and objective measures of what it reasonable or appropriate. Of course, empirically informed regulation cannot produce a perfect set of legal rules but intends to provide institutionalised monitoring and feedback of its own utility. In order for this goal to be achieved, a change in the conception of the law-making process is necessary that can accommodate the rapid pace of technological advances.

Rather than insisting on creating sustainable and static laws that should be revised after a period of twenty years, the rapid technological progress calls for "a post-bureaucratic vision" of law, in which regulators govern through the use of auditing and continuous adaptation to diverse and changing environments (Simon, 2015).

The 'post-bureaucratic' or performance-based approach has emerged in the private sector as industries have sought flexibility to adapt to more volatile economic circumstances and, as a consequence, an analogous need for governments has been created to respond to fluidity and diversity by importing elements of the post-bureaucratic view to the public sector. The differences in the structure of the bureaucratic and post-bureaucratic approaches also reflect differences in the organisational premises (Sabel & Simon, 2011). Contrary to the former approach that tends to presuppose bureaucratic organisation, the latter often arises from performance-based organisation. According to the bureaucratic approach, or else canonical doctrine, which was the dominant paradigm of efficient large-scale organisation in the past, authority depends on previously chosen values and goals, while rules are relatively inflexible and difficult to change. In the bureaucratic view, the rule is the most important type of norm. The canonical doctrine also endorses the reactive approach to error detection. By contrast, in post-bureaucratic organisation, legitimacy depends less on prior authorisation and more on transparency, whereas the key type of norm is the plan instead of the rule. Even though plans may contain rules, they are regarded more comprehensive and provisional, and less categorically prescriptive and tight compared to bureaucratic rules. Plans typically set out procedures for monitoring their own implementation and for frequent assessment in the light of information deriving from monitoring. Another significant element introduced in the post-bureaucratic organisation is the adoption of a proactive approach to error detection relying on audits aimed at ex ante preventing the undesirable consequences rather than ex post mitigating them (Simon, 2015). The differences between the two approaches indicate that the need for updating the canon, which stems from the need for governments to adapt to the rapidly changing circumstances, requires a thorough reconsideration of the bureaucratic doctrine that lies in broadening its focus and altering its organisational structure.

7.3.6. Incorporating adaptive management process into cyber security regulation

The main characteristic of the concept of adaptive management as a learning-based decision process is that it promotes flexible decision making that can be adjusted in the face of uncertainties. Adaptive management serves as a means of adjusting policies as part of an

iterative process following careful monitoring of the outcomes of certain actions. This does not mean that it should be conceived as a random 'trial and error' process; instead, emphasis should be placed on iterative learning, which aims to reduce uncertainty, and on improved management as a result of learning (Allan, 2007). In other words, this process can be pictured as a 'feedback loop'[80] of monitoring, evaluation, and management adjustments that focuses specifically on learning about the impacts of management. Within each iteration of the overall cycle, multiple iteration of this loop may occur. Adaptive management is designed to improve understanding of how a system works in order to achieve the desired outcomes. To this end, the management process involves formulating the resource problem, developing conceptual models, and identifying actions that might be used to resolve the problem (Bond, 2016). Through the monitoring of the outcomes of each operation, actual outcomes are compared against predicted outcomes, with the comparative results fed back into decision making to produce more effective decision making. A significant prerequisite for the effective application of the process of adaptive management is the involvement of organisations or individuals who use, influence, or have an interest in a given resource, early in the adaptive management cycle so as to help assess the problem, design activities to solve it, implement and monitor those activities, and participate in the evaluation of results (Williams & Brown, 2014).

7.3.6.1. Smart Governance

The concept of Smart Governance introduces the idea of creating a system where the regulated entities are subject to real-time measurement to better achieve the regulatory goals. Smart Governance is dynamic in that it involves interactive loops dependent on each other and constantly readjusted according to the feedback provided. The aim of the Smart Governance model lies in helping regulators make better decisions on how to improve the regulatory process and achieve the goals set out in their mandate. The framework introduced by this model allows regulators to adopt an approach to problem-solving similar to that employed by companies since regulators can utilise it in order to design public policies, improve the implementation of regulation and work together when updating legislation

[80] *Feedback loop* is a term commonly used in the field of Economics to refer to a situation where part of the output of a process is used for new input, *See* Koteshwar Chirumalla, 2017, Clarifying the feedback loop concept for innovation capability: A literature review, Presented at The XXVIII ISPIM Innovation Conference – Composing the Innovation Symphony, Austria, Vienna on 18-21 June 2017, available at: https://www.researchgate.net/publication/316857374_Clarifying_the_feedback_loop_concept_for_innovation_capability_A_literature_review.

(Cohen, 2016). The application of this model to the regulatory process could be achieved using the following three methods: the Regulatory Management Method, the Regulatory Auditor Method, and the Regulatory Oversight Method. The first method involves the collaboration of the regulator with the relevant advisory committee in order to identify relevant pieces of dynamic performance data that can be collected from all regulated entities and applied to a particular outcome-based goal. In this scenario, the regulator is responsible for managing the entire process as he is in charge of housing databases with the information collected, using modern analytical techniques for deriving data insights from the data, and using his enforcement authority against entities falling short of achieving specific results. In the context of the second method, the onus of responsibility is placed on the regulated entities, which are in charge of maintaining databases, creating algorithms to analyse the data, demonstrating that they are meeting the goals defined, and reporting their progress towards achieving those goals on a regular basis. Although the regulator still sets out a series of dynamic metrics for regulated entities and can use his authority in case the results of the report fall below a certain threshold, the fact that the regulated entities are those which determine the means for best achieving the regulatory goals allows them the flexibility to innovate. The same holds true as regards the last method, where the only difference compared to the second one is that the regulator is also assigned to play an oversight role, since the regulated entities ought to create an internal independent auditor, who is subject to controls by the regulator; at the same time, the regulator is still responsible for subjecting the regulated entities to enforcement if they fail to utilise the controls set out by him. The internal auditor is also subject to annual review of his programme (Cohen, 2016). Relevant to the notion of auditor seems to be the concept of 'algorithmist' introduced by Mayer-Schönberger and Cukier (2013), who envision a new type of profession akin to that of auditor that would ensure accountability, traceability and confidence in Big data predictions in the same way auditors did with financial information in the early 20th century. In particular, the authors describe two types of algorithmists: the internal algorithmist who would be a sort of Big data ombudsman, and the external algorithmist who would consult with government on how to best use Big data in the public sector.

7.3.6.2. The Boyd cycle in the regulation-making context

The so-called O.O.D.A loop, or else Boyd cycle, is an information management concept and refers to a decision-making model currently being applied in business and technology contexts (Boyd, 1987); it is commonly used as a tool for developing web applications as it

allows developers to identify the changes needed and release improvements to the application by experimenting on new things (Hammersely, 2012). The Boyd cycle comprises the four stages of *observing* (by gathering inputs from the environment), *orienting* (by creating a model of the situational reality based on the data collected), *deciding* (by using the knowledge acquired as a new basis), and *acting* (by translating this knowledge into action) (Richards, 2012). Mark Fell (2013) advocates the adoption of the Boyd cycle, also referred to as Smart Governance cycle, in the field of policymaking in order to enable governments to transit to a more agile, collaborative and insightful regulatory model. Applying the Boyd cycle to regulation is a process consisting of the following steps, each of which should be exercised repeatedly to gain significant benefits and ensure accountability. The first step would require regulatory bodies to collect relevant data from all regulated actors regarding their performance. Unlike the practices currently used by regulatory bodies, which already collect massive amounts of data sometimes irrelevant or in an inefficient manner, they could do so through digitisation by using an Application Programming Interface (API) that regulated entities would plug into to submit relevant data. Two crucial elements of this process are collaboration and harmonisation. In determining the relevant data points for a particular regulated environment, experts from both sides of regulators and regulated should look at specific pieces of information that all actors in a regulated environment should have and then harmonise the data request across them. Not only does comparison between regulated actors become easier when data points are harmonised, but also a comprehensive dataset can be created for regulators to work on.

The second step, which is the most significant part of the cycle, is about making sense of the data collected and transforming these data into information in order to choose the best course of action. Encouraging the collection of data from disparate sources and organising them in an understandable manner would enable meaningful insights to be derived which, in turn, would effectively support the decision-making process (Australian Government, 2013). At the third stage of the cycle, experts are in charge of both creating and applying algorithms to gain insights from the database. In this context, the notion of algorithm should not be considered in its purely technical sense but rather as a set of rules to be followed in an operation (Cohen, 2016). Using data insights regulators would be able to think broader about the changes needed to be introduced into a system and the ways of doing so in order to achieve the set regulatory goals. It is essential that the three aforementioned steps are constantly reviewed and readjusted to the rapidly changing environment based on the

feedback generated from previous iterations. The role of feedback loops lies in guiding regulators to, for example, reform the system in case databases are not structured in an efficient manner, or readjust the calculations if algorithms are not leading to meaningful insights. Even though regulators are likely to encounter difficulties in putting in the right place feedback loops that allow for corrections and innovation, it is still significant regulators to be agile in terms of both their processes and means of achieving their objectives. Finally, the last step of the Boyd cycle includes putting the right inputs into action, that is, interpreting the insights in context and deciding how to implement them into regulation (Cohen, 2016).

In order for the Boyd cycle to effectively operate in a regulation-making context, it is important that the right actors – referred to as "intervention agents" – also constitute part of the entire process (Fell, 2013). In the context of regulatory decision making, appropriate intervention agents include computer algorithms, crowds and recognised experts. Good innovation in policymaking involves finding the 'intervention mix' suitable to address a particular problem, and then evolve that mix depending on the stage of the Boyd cycle. The inclusion of intervention agents and, especially, the collaboration of technical and policy experts from government, industry, academia, non-governmental and consumer groups (market wide consultation) in the early stages of the decision-making process, and throughout the process, can ensure that regulators are not left behind from technical developments in industry since it helps them overcome their lack of technical knowledge, a problem that plagued traditional performance standards. Attention should be paid to the process of selecting the right actors, which should by no means be self-selecting as this would result in bias and expertise risk. In areas where a combination of knowledge and initiative is required, for example, computer algorithms may lead to pure rule-based decisions but it should also be considered that they lack the sensitivity to context held by humans.

7.3.6.3. Smart Governance & dynamic performance standards: A new model for developing cyber security regulation

In a world of pervasive rapid development, performance standards appear to dominate over design standards as most modern businesses use some form of performance standards. Among the various benefits of performance standards, the most significant ones that make them prevail are the ability to directly address the regulatory goal and account for changes in the practices for regulated entities, and the fact that they empower innovation in compliance methods by allowing for discretion in implementation while incentivising developments

occurring in industry (Bensanko, 1987; Coglianese et al., 2003). Nevertheless, despite the fact that traditional performance standards provide the regulated entities with more flexibility compared to design-based regulation, they are still unable to iterate and thus there exists the danger of locking a particular practice among industry ending up playing the same role with design standards (Stewart, 1981). Moreover, enforceability issues often arise due to regulators' difficulty in monitoring the compliance of entities subject to performance standards and thus determining whether regulatory objectives have been achieved (Breyer, 1982). The recently developed concept of 'dynamic performance standards' aspires to overcome the likely stagnancy of traditional performance standards by trying to avoid the pitfalls of the latter. First, this concept addresses the problem of assessing whether a goal is being met, which has resulted in an 'information gap' between industry and regulators, by utilising data analytics techniques to make easier the process of measuring and analysing the performance of the regulated entities. Second, the iterative nature of dynamic performance standards, which enables regulators to innovate at a similar pace as industry, aims at mitigating the regulatory risk involved in the static nature of classical performance standards; the iteration process is based upon new data and new insights. Third, the lack of regulators' technical knowledge to measure, monitor and iterate the standard has led to difficulties in creating effective performance standards, a goal that can be achieved only with the collaboration of all concerned stakeholders – a feature introduced by the concept of dynamic performance standards (Bianchi, 2016).

As stated in the World Economic Forum in 2012, "[we] are living in the most complex, interdependent and fast-paced era in recent history". As a matter of fact, technology is nowadays the key driver and dramatic changes in the field of technology are enabling business models that were not possible twenty years ago. However, the innovations in the industry sector operate within a regulatory system that is struggling to keep pace as the transformation taking place is challenging the existing regulatory approach. Instead of the law continuously trying to keep pace with technology, it would be better to embrace technological advances and benefit from the same cutting-edge practices that are revolutionising industry. The key challenge is not only to set out the key objectives of a regulation, which is undoubtedly a crucial part of the entire process, but mainly to define and operate regulation able to realise the governmental goals set in the most effective and efficient way possible. In this regard, the existing regulatory approach to technology-related matters is falling short of achieving the goals initially defined. In order for regulation to

innovate and transform in a manner akin to the way industry evolves, a shift in the regulatory decision-making model is required, which will be able to meet the needs of the current complex and fast-paced business environment. The new model for developing regulation would be based upon the concept of dynamic performance standards for measurement and the concept of Smart Governance for implementation and monitoring, which suggests the integration of practices, such as gathering market data and using data analytic tools, in the regulatory system to create a better-informed regulatory development process. By combining the use of technology and data with a collaborative and iterative process to measure the performance of regulated entities, regulators will be able to obtain unique insights and better deliver the goals underlying regulation thus keeping pace with the highly innovative and transforming industry.

The Big data revolution, which came as a result of increases in data acquisition capability, data storage, computing power and algorithmic design, has enabled better insights in developing technology. Big data used in the technology sector has transformed not merely the way traditional industry conducts business but also the way governments provide their services and thus improving governments' efficiency in terms of procurement and law enforcement issues. By contrast, Big data has so far failed to reform the process by which regulation is created and implemented. Introducing, however, the data analytics element to regulation may enhance regulator's ability to measure and analyse performance standards and possibly lead to a transformation of the regulatory process of designing, implementing and improving policy and legislation in collaboration with stakeholders. In a data dominated society characterised by fast-paced transformation, where our basic understanding of how to make decisions and comprehend reality is being challenged, creating a public policy process requires a particular mindset – the so-called "big data mindset".

7.3.6.4. Incorporating smart regulation principles into cyber security regulation

Smart regulation has various facets, some of which will be analysed below in the data protection context. To begin with, its primary characteristic lies in the fact that it encourages the adoption of an incentive-based approach to compliance that takes into consideration the complex and diverse compliance motivations of each organisation. Such motivations could potentially be provided in the form of reduction in liability, defences in data breach litigation or even reputational payoffs. It is crucial that stakeholders learn from industries that have already applied incentive-based mechanisms so that any unintended consequences of such

incentives can be minimised to the extent possible. Recognising that companies often have plural and overlapping motivations (e.g. preventing brand and reputational damage, boosting the confidence of customers and business partners) and taking into account their business drivers so that they are allowed to innovate is another element of smart regulation. Furthermore, smart regulation encourages the adoption of an "open culture" where the compliance of organisations with the applicable data protection laws will transparently be improved through guidance and support instead of sanctions. It is important that data protection authorities dialogue productively with organisations in order to find mutually acceptable solutions (Vranaki et al., 2016).

Regulatory pluralism constitutes one of the primary aspects of smart regulation, which endorses the idea that the recruitment of third parties in the regulatory process may provide for improved outcomes (Klingbeil, 2010)). Contrary to the traditional regulatory process that is restricted to government considering it as an omnipotent form of regulatory authority – a notion of government that tends to become outmoded – smart regulation aspires to introduce a model of governance, where a range of third parties, both commercial and non-commercial, will be given the chance to play a valuable role in the regulatory process, acting as quasi-regulators. As a matter of fact, in some instances, quasi-regulation might be more potent than government intervention, especially in areas of commercial activity which impact on the performance of industry and thus render direct government intervention impractical. Moreover, given that "smart regulation is not about more or less regulation, it is about delivering results in the least burdensome way" (European Commission, 2017), the involvement of third parties in the regulatory process can assist in taking the weight off government intervention and hence allow government to reserve its resources for situations where direct intervention is the sole viable regulatory alternative. An additional benefit of smart regulation is that empowering participants who are in the best position to act as 'surrogate regulators' can generate a regulatory outcome with the potential to be mutually reinforcing due to the multiplicity of regulatory signals. In terms of the mechanisms through which government should seek to engage third parties in the regulatory process, the provision of adequate information can serve as a starting point; reliable data on the performance of industrial firms would enable those actors who are in a position to exert influence (e.g. in the commercial sphere) to make objective judgments with regard to the issues at stake (Gunningham & Sinclair, 1998). More importantly, the role of government must principally be that of a catalyst or a facilitator which enables a coordinated collaboration between

surrogate regulators by filling any gaps that may exist and facilitating links between the different layers of the regulatory process. Instead of trying to find ways to engage third parties in direct intervention, it would be preferable for government to create the necessary preconditions for them to assume a greater share of the regulatory burden.

In its 2010 Communication on Smart Regulation in the European Union, the European Commission defines the three main characteristics of smart regulation as follows: first, the concept of smart regulation should be integrated throughout the policy and law-making cycle – from the design of a piece of legislation to implementation, enforcement, evaluation and revision; second, the regulatory process must be a shared responsibility of the European institutions and Member States; third, the views of those most affected by regulation can play a key role in smart regulation. In terms of the participation of stakeholders in the regulatory process, the Commission describes the inputs on impact assessments from citizens and stakeholders as an essential element of smart regulation in order to deliver good quality policy proposals and stresses that the implementation of smart regulation objectives requires consulting them both when developing policies and when evaluating whether the regulatory outcome is in accordance with the initially defined regulatory goals. To this end, the Commission has established a platform called 'REFIT Platform', where any stakeholder can present their views on the impact of EU law and also make suggestions on how the legislation can be improved. In particular, interested citizens and stakeholders can contribute by sharing their opinions on forthcoming initiatives outlined by the Commission (roadmaps), legislative proposals and their economic and social impact, or by evaluating existing policies (European Commission, 2015). The role of this platform is to provide a basis for inclusive work on a common agenda involving civil society, social partners and high level experts from business appointed through an open and transparent process. Open and transparent decision-making process is considered to be of vital importance for effective regulation and thus all major initiatives should be subject to thorough assessment. Two additional principles of effective regulation, as set out by the Commission, dictate that any regulatory action needs to be based on evidence and adequate information (e.g. detailed information about roadmaps should be available online), while regulatory burden on businesses, citizens or public administrations must be kept to a minimum. Furthermore, good practice should be extended to the entire policy cycle, which entails ex post evaluation of regulatory performance, assessment of economic and social impacts, as well as understanding the full benefits of regulation and not only focusing on the costs of its implementation (European Commission, 2017).

7.4. Conclusion

As demonstrated above, performance-based regulation may also generate undesirable side effects in case too much discretion is granted to the regulated entities. For instance, if regulation is subject to wide interpretation as to the required level of security, it is more likely that a company opts for lax security measures, a choice that may be beneficial for the company in the short run but will increase the likelihood of security breaches in the long term. In determining whether to use a performance standard, and if so, what specific type of standard to adopt, given that performance standards themselves differ in their specificity, measurability and feasibility, regulators need to make sure that the necessary conditions under which a performance standard should be applied are being met. The above analysis of the considerations regarding the weaknesses of the performance-based regulation, coupled with the difficulties associated with implementing such regulation, indicates that it is not necessarily a 'silver bullet' regulatory strategy that can be applicable under any circumstances. Some situations call for a hybrid approach that helps minimise the limitations of performance-based regulation. A hybrid approach is a blend of regulatory instruments that may take various forms. One such a form is the example of tiered regulation, which mandates performance thresholds and also provides prescriptive guidance suggesting the use of specific technologies. Another type of hybrid approach is a regulation that requires specific technological designs but also includes equivalency clauses or provisions for alternative enforcement mechanisms; provided that the regulated entities can achieve a comparable level of performance through other means, they are allowed to opt out of the prescriptive standards (Coglianese et al., 2003). The main challenge for regulators is to obtain a thorough understanding of the nature of the problem that calls for government intervention, which will enable them to identify the conditions under which different regulatory instruments are appropriate, while also keeping an eye on changing conditions or new alternatives (Daly & Foushee, 1981). It is crucial that a regulator's decision is independent of political economy issues. Imposing strict security obligations on companies collecting personal data is likely to raise objections on the part of the private sector, but restricting law-making to regulations that companies do not oppose allows private organisations rather than citizens to govern, an outcome that is incompatible with the principles of democracy.

It should be recognised that the regulatory environment differs from that of a firm in many aspects, and therefore it would be politically and socially unacceptable to experiment on a

market-wide scale, where a mistake is magnified many times over. Yet, this does not mean that policymakers should abstain from adopting new approaches to regulation but, instead, they should find a way of doing so that increases market confidence and the quality of intervention by using, for instance, firm level pilots or market simulations while creating safe harbours which would ensure that regulated entities have confidence in the pilots and simulations. In order for regulation to remain grounded in the realm of technical feasibility, it is essential that technological advances such as data analytics techniques are actually implemented in the regulatory structures. Based on the premise that "a 21rst-century agency should use 21rst-century tools", the techniques offered by the concepts of Dynamic Performance Standards, in combination with the Smart Governance model, can be utilised by policymakers and regulators to glean new insights and hence revolutionise the regulatory decision-making process. When the risk-based approach to regulation, which requires that legislators and regulators focus on those areas that present the greatest risk to the regulatory objectives, is applied through the lens of the aforementioned concepts, is likely to not only allow for better results but also release useful resources for public authorities to concentrate their enforcement activities where it really matters. While respecting the role of regulators, the new approach suggested supplements it with cutting-edge thinking that involves decision making based upon iterative data analysis, collaboration with the most appropriate actors on an issue-by-issue basis and focus placed on performance rather than design. In this way, regulators can benefit from the opportunity to better deliver policy goals while enabling new business and operating models. This paradigm shift, however, requires a change in terms of both methodology and way of thinking. In particular, all the elements that form the regulatory decision-making process need to be reconsidered and take differently shaped dimensions. Therefore, a new regulatory model would require understanding different business models and emerging technologies, involvement of new and more stakeholders, monitoring of market mechanisms and practices, and utilisation of new tools such as data management and algorithmic solutions. More importantly, a change in culture is necessary to embrace the new regulatory model that challenges the traditional approach to regulation. It is noteworthy that the Smart Governance model places emphasis on the question of how to achieve better regulation in the place of the question of whether more or less regulation is needed.

CHAPTER 8: Need for an Alternative Regulatory Approach to Data Security

8.1. Introduction

This chapter aims to build a regulatory model able to address the issues arising from the problematic nature of the identifiability concept in the current EU legal framework surrounding the security obligations of data controllers. To this end, the first part examines the benefits of the risk-based approach to data protection and the notion of risk as it appears in the provisions of the GDPR, according to which data processing operations which involve high risk trigger additional compliance obligations. The notion of risk under the GDPR is linked to the notions of threat and harm as the extent of the threat and the likelihood and gravity that different processing operations raise to individuals' rights and freedoms play a crucial role in determining whether a certain processing activity is of high risk. The second part of the chapter provides insights on the need for developing a regulatory model able to overcome the challenges posed by the binary distinction between personal and anonymous data in the post-anonymisation era, which will provide the necessary incentives for companies to implement adequate security mechanisms. The analysis of the EU legislator's attempt to address the utility-privacy trade-off by introducing the pseudonymisation concept indicates the shortcomings of this approach that seems to raise a range of technical and legal issues. The chapter proceeds with the debate over the de-identification of personal data in order to highlight the need for introducing a spectrum of identifiability degrees, which recognises that the categories of identified and anonymous data are not the only possible states of data but rather the two end points of the identifiability spectrum. Drawing on the literature on the degrees of data identifiability, a regulatory model is proposed that aspires to overcome the issues stemming from both the de-identification debate and the technology neutral approach to the security obligations of data controllers, by systematically tailoring the security requirements depending on the state of the data, which is contingent upon the level of the re-identification risk resulting from the interplay between certain risk factors and varies depending on the data environment. The model aims to provide clarity to data controllers' security obligations by laying down high-level security objectives and recommending

examples of security measures to be implemented according to the security level required in each case.

8.2. Need for a Risk-based Approach to the Security Obligations of Data Controllers

Developments in Open and Big data result in data being processed on a scale never achieved before due to the incredibly fast computer processing speeds. Besides the benefits that such technological advances entail, immense threats to privacy have also emerged, which are driving an unprecedented need for reliable anonymisation techniques. As demonstrated above, although anonymisation had for long been viewed as the solution to unlock the broader utility present in massive datasets, it is now getting more and more difficult to carry out the anonymisation process with confidence. However, third parties are still allowed to process anonymous data lawfully since any link relating to the data subjects is considered to have been removed. It is crucial, though, that even in this case third parties take into account the (even low) possibility of being able to identify individuals, at which point data becomes personal and thus falls within the ambit of the data protection laws. The challenges posed by the failure to achieve perfect anonymisation herald the need for a new data protection paradigm. Instead of dwelling on the future fragility of the technique of anonymisation as a solution for protecting personal data, which appears to be increasingly irrelevant in the face of widespread sophisticated profiling techniques, it would be more realistic to build a new model whose focal point will be the risk and the effects of the processing of personal data on the data subjects (Gratton, 2014). In the post-anonymisation era, the risk of re-identification can serve as a point of reference for regulators to set a threshold below which personal data can be considered protected under given circumstances and thus treated accordingly.

As the pace of technological change outstrips the conventional thinking of regulators and businesses, a calibrated and risk-based approach to data protection might improve the ability of businesses to take a better-informed and better-structured approach to the processing of the massive personal information they collect, store, use and share on a daily basis. The principal advantage of the risk-based approach lies in the fact that it goes beyond mere compliance with regulatory requirements and allows organisations to find ways of not only implementing

data protection requirements on the ground but also ways of demonstrating their compliance (Hunton & Williams LLP, 2014). In doing so, the risk-based approach helps clarify and communicate the underlying rationale of data protection regime. In contrast to regulatory approaches that merely focus on inputs and classifications, regulatory approaches based on risks and outcomes aim at securing the object of data protection laws as they examine the overall outcome of the processing of a particular piece of information, by the data controller in question, in terms of the risk to individuals' fundamental rights and freedoms (Article 29 WP, 2014b). Such a contextualised approach that hinges upon the impact on individuals can lead to increased standards of protection by ensuring that processing operations resulting in privacy harm are brought within the scope of the data protection regime, irrespective of their exact nature or form, and by incentivising data controllers to invest in activities aimed at benefitting individual privacy (Brown, 2010).

The increased focus on the notion of risk as a touchstone for data protection regulation has been one of the most significant developments in the data protection field over the last decade. The risk-based approach adopted by both the DPD and GDPR is based on the premise that organisations handling personal data must devote more resources to the processing operations that raise the most serious threats. Given that risk varies across industries, instead of imposing one-size-fits-all approach, the law promotes a nuanced approach according to which organisations adjust protection measures to their business model (Schwartz, 2016). Even though the concept of risk is central, there is no agreed definition. Instead, it is often used flexibly to apply to different components of 'risk'; for instance, one might refer to the risk of loss of confidentiality or financial loss as the results of a data breach, and at the same time, refer to the risk of a data breach, which does not necessarily lead to a harm or damage. In the risk assessment context, risk is conceived as a scenario which combines a feared situation (e.g. breach of the data processing security and the consequences the breach entails) and the possibilities that the feared situation will take place (e.g. threats to the supporting assets); in this regard, risk level is estimated in terms of gravity (i.e. the degree of its impact) and likelihood (i.e. the probability of occurrence) (CNIL, 2010).

In the data protection context, the most broadly used definition of "risk" encompasses the concept of "unacceptable risk", that is, a threat to, or a loss of a valued outcome that cannot be mitigated through the implementation of effective controls and is disproportionate to the expected benefits. Therefore, "risk" equals the probability that a data processing activity will result in such an impact (Hunton & Williams LLP, 2016). The concept of "unacceptable risk"

is also associated with the risk-tolerant nature of any data protection regime, whose aim is not to eliminate any possible risks but rather identify and eliminate the unacceptable risks. In other words, the data protection measures that should be applied by data controllers should be designed to reduce risk as much as reasonable and practicable in the light of the available protection mechanisms, the costs and efforts required as well as the intended benefits. Although there is no concrete definition of 'risk' in the GDPR, the Regulation provides interpretative guidance on what may constitute risk and harm. To begin with, the GDPR seems to focus on one type of risk, that is, the adverse risk of processing activities to the individual. In particular, Recital 74 unambiguously states that data controllers should implement "appropriate and effective measures" in light of the risks to the rights and freedoms of the data subjects generated by the data processing. The risk-based approach and the notion of scalability, which envisages that the required compliance measures should also take into account the nature, scope, context and purposes of the processing operations, are closely linked mechanisms incentivising accountability. Despite the fact that solely the risks to individuals are explicitly mentioned in the GDPR, it is noteworthy that such risks are closely linked to other aspects of risk, because organisations usually incorporate the individual-based risk assessments into their broader enterprise risk management systems, which actually evaluate risks to the organisation, such as financial, reputational, litigation risks, or corporate opportunity risks associated with the organisation's business and profit objectives. As a matter of fact, guidance on data protection impact assessments (DPIAs) and existing practices suggest that risk assessments under a DPIA may go beyond assessing the risks to individuals, but may assess the risks of a wider set of stakeholders, including the risk to the organisation itself or to society at large (ICO, 2014b).

Recital 75 of the GDPR prescribes that the actual risk level should be determined by considering the "likelihood" and "severity" of the risk to the rights and freedoms of natural persons. The term "likelihood" refers to the possibility of a risk or its impact to materialise, while "severity" means the magnitude of the risk or its impact in case it materialises. Thus, organisations are obliged to assess the likelihood and severity of their personal data processing activities and apply processes that enable them to reliably and consistently weigh the risks and harms against these criteria. This requirement does not mean that the protection of the rights of the data subjects depends on the level of risk of the processing in question, since individual rights apply in full regardless of the risk level. But it is organisations' responsibility to modulate their data protection compliance according to the level of risk their

processing operations pose to the individuals' rights and freedoms. Even though there is no consensus on the specific criteria to be considered when evaluating the likelihood and the severity of a risk, other than the nature, scope, context and purposes of processing (Recital 76), three risk elements can be identified in Recital 75 – risky processing, potential threats and harms. Therefore, in determining whether there appears a significant risk to the rights and freedoms of the data subjects, the question that needs to be answered is whether there is a significant possibility that the particular threat could lead to the particular harm with a significant degree of seriousness. In terms of the first risk assessment element, Recital 75 provides examples of processing activities that qualify as potentially risky, such as the processing of a large amount of data affecting a large number of individuals, the processing of special categories of data, the processing of the personal data of vulnerable natural persons etc.; this list of risky operations is not exhaustive and hence organisations ought to consider other possible risk elements in each case based on context. When considering the potential threats in a risk-based environment, particular attention must be placed on the use of information, which poses the greatest threat. In this respect, as threats arising from data processing can be regarded threats that are likely to appear over the lifecycle of data, such as the excessive collection of data, the use or storage of inaccurate or outdated data, and the inappropriate use of data. As with the risky processing activities, the assessment of potential threats should also be contextual, which means that context needs to be recognised as an important factor in determining the level of threat and its potential to cause harm. Finally, Recital 75 identifies two types of harms that potentially risky processing activities or threats might present, that is, material or non-material damage. The first category of damage can be conceived as any tangible, physical or economic harm to the individuals, such as financial loss, bodily harm, loss of freedom of movement, whereas the second category refers to the intangible distress caused to individuals which may entail unacceptable intrusion into private life, deprivation of control over personal data, reputational harm etc.

It is worth mentioning that the GDPR takes two different approaches to the concept of risk. On the one hand, it conceives risk sequentially, which means that data controllers are obliged to adopt stronger data protection measures as their data processing poses increased possibilities of harm. On the other hand, in contrast to a threat model based on a continuum, the GDPR sometimes divides risk into two mutually exclusive categories – risk and high risk (Recital 76). The significance of this distinction lies in the fact that data processing operations which involve high risk trigger additional compliance obligations and thus the

identification of the amount of risk and the question of which category a certain operation falls into are matters of crucial importance. Before discussing the second approach, the specific obligations calibrated by risk will be briefly analysed. In comparison with the DPD, the GDPR appears to broaden the relevance of risk as it is not only explicitly based on the notion of a risk-based approach but it also seeks to tailor data controllers' and processors' legal obligations to the risks that the data processing presents to a far greater extent than the DPD. Contrary to the DPD, which makes only sparing use of the risk concept for its security requirements (Article 17 requires that data controllers "ensure a level of security appropriate to the risks represented by the processing and the nature of the data to be protected"), the GDPR also uses this concept as a cornerstone in order to impose accountability obligations on both data controllers and processors.

For instance, Article 24 contains a general accountability provision calling for data controllers to implement technical and organisational measures commensurate to the identified threat after assessing the "risks of varying likelihood and severity for the rights and freedoms of the individuals". This requirement means that a controller must build, implement and be able to demonstrate a data protection program, where the protection measures will be more intensive in respect of processing that creates risks; the actual compliance programs may vary between controllers depending on the various levels of risks associated with their processing. The risk principle is also reflected in Article 25 which requires data controllers to take the risk factor into account when determining the building-in of data protection principles. In another use of the risk principle, Article 30 absolves an enterprise or an organisation with fewer than 250 employees from the obligation to maintain a record of processing activities as long as its activities are not likely to result in "a risk to the rights and freedoms of data subjects"; in this context, risk provides a threshold below which data controllers are exempt from an obligation. Likewise, data controllers are not obliged to notify a data breach to the supervisory authority when the breach is "unlikely to result in a risk to the rights and freedoms of natural persons" (Article 33). Finally, the GDPR uses risk to organise its security principle in the same way as the DPD; Article 32 of the GDPR is in line with and elaborates Article 17 of the DPD. Pursuant to Article 32, data controllers and processors should implement "appropriate technical and organisational measures to ensure a level of security appropriate to the risk" presented by their processing activities. In determining the appropriate level of security, Article 32 calls for heightened attention to a special group of security risks, the risks that may arise from "accidental or unlawful

destruction, loss, alteration, unauthorised disclosure of, or access to personal data transmitted, stored or otherwise processed."

As far as the second approach adopted by the GDPR is concerned, the extent of the threat that different processing operations raise to individuals' rights and freedoms plays a crucial role in determining whether a certain processing activity will be regarded as involving high risk and thus will trigger additional compliance obligations, or it will be seen as raising lower risks to the fundamental rights and freedoms of individuals and, as a consequence, will result in fewer compliance obligations. It is worth mentioning that the high-risk category was absent from the European Commission's proposal for a new data protection regime of 25 January 2012, whilst in the amendments of 12 March 2014 suggested by the European Parliament, this term was introduced in two different procedural stages but none of them survived subsequent negotiations. The 'high-risk' concept was eventually established by the European Council in its draft of 11 June 2015, which resulted in its integration in the final text of the GDPR. To further clarify the meaning of high-risk processing, the GDPR provides guidance on what may constitute high-risk processing by giving examples of the characteristics of potentially high-risk processing operations, which should, however, be viewed by organisations as being part, not the only features, of a risk assessment process. Generally speaking, Recital 76 suggests that an objective assessment is required for the assessment of risk, which would determine "whether data processing operations involve a risk or a high risk". More specifically, Recital 89 states that types of processing likely to result in high risk include processing activities where new technologies are used, activities that "are of a new kind and where no data protection impact assessment has been carried out before by the controller", as well as activities that "become necessary in the light of the time that has elapsed since the initial processing". Article 35(1) identifies the use of new technologies during the processing as a possible circumstance of 'high risk' requiring a data protection impact assessment, while Article 35(3) appears to establish the following 'default' high-risk categories: the "systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person"; the "processing on a large scale" of special categories of data and of personal data relating to criminal convictions and offences; and the "systematic monitoring of a publicly accessible area on a large scale". The elements of these potentially risky processing operations should, by no means, be considered as the sole triggers for high-

risk status, but must be coupled with additional high-risk characteristics based on nature, context, scope and purpose of processing. The GDPR imposes specific obligations on data controllers that are triggered only in the cases of high-risk processing. The leading example of such an obligation concerns the data breach notification requirement, according to which controllers ought to notify not only the responsible supervisory authority (Article 33), but also disclose the breach to the data subject without undue delay when the breach "is likely to result in a high risk to the rights and freedoms of natural persons" (Article 34(1)). In addition to the requirement for conducting a data protection impact assessment in cases of high risk processing, as described above, the GDPR obliges controllers to also seek formal advice from the responsible supervisory authority before processing personal data in case the impact assessment "indicates that the processing would result in a high risk in the absence of measures taken by the controller to mitigate the risk" (Article 36(1)).

8.3. Need for a Regulatory Model Imposing Nuanced Security Obligations

8.3.1. Need for a regulatory model able to effectively address the trade-off between data utility and privacy

The need for addressing the utility-privacy trade-off is mostly evident in a wide range of industries and research fields, where implementing de-identification techniques in order to protect personal data is likely to render the data useless for other purposes. Computational privacy protection mechanisms inherently involve a trade-off between utility and privacy since the data elements have to be somehow distorted in order to safeguard privacy (Wu, 2013). This becomes apparent when deploying, for example, the technique of differential privacy, which involves adding noise to the data, or the technique of k-anonymity, which is often achieved by generalising and suppressing data contained in a database (Sweeney, 2002). If the variables of interest are stripped off, then their utility diminishes especially for researchers and data analysts. In areas such as healthcare or research, for instance, where the ability to re-link information to the data subject is critically important, striving for irreversible anonymisation may not only degrade the quality of data but also impede the utility of essential services. In other words, higher standards for de-identification may result in lower-

value de-identified data, which will have a chilling effect on the motivations of data controllers.

In the case of geolocation and traffic applications, for example, location services must track the user's geolocation at particular points in time in order to provide real-time traffic information. When such data is sent to a phone carrier, operating system or location service provider, any identifiers are usually hashed and the hashed traffic data are then placed in a data vault and are enriched with additional location data returned to the user's device. The processing of payment transactions is another example indicative of the values of data processing both to merchants and consumers. Not only can the information collected be used to improve operational efficiencies, but also to prevent fraud and detect theft. Often payment data needs to be quickly linkable to an individual so that the purchaser's identity can be readily confirmed (Polonetsky et al., 2016). In order to maintain unique payment records without identifying users, payment processing companies remove any personally identifiable data from the transaction data and subject the account numbers to a one-way hash. As far as non-transaction data is concerned, it is aggregated and reviewed before being combined with de-identified transaction data, and then information is placed into separate warehouses, where it can be further aggregated into larger datasets depending on the sensitivity of data. Such aggregated data can then be accessed by data analysts or merchants, and can be combined with other macroeconomic data to gain further insights.

Therefore, striking the appropriate balance between data utility and data protection is the only means of providing those in charge of processing personal data with the necessary incentives to apply practical security measures that would also allow them offer significant services based on de-identified data. The core of the protection-utility trade-off lies in finding ways to structure data so that meaningful information can be derived from it, while preventing its unintended use, fraud etc. Even though data protection and data utility may appear to be two mutually exclusive notions, at first glance, the combination of technical, administrative and legal controls has the potential to effectively address this trade-off. In the abovementioned example of geolocation data, such safeguards would include hashing individual identifiers or aggregating data after a certain period of time and applying contractual use restrictions to protect location data, while in the case of traffic data, before hashed traffic data being shared with third parties for research or marketing purposes, it should be further aggregated into traffic reports and the data practices of data recipients should be reviewed.

A study conducted by a group of experts at Harvard University coming from different disciplines (computer science, social science, statistics and law) examines ways to develop computational tools for measuring privacy and data utility in order to achieve an optimal privacy-utility trade-off. The group adopted a holistic approach by complementing these tools with a set of legal instruments for social scientists to use when dealing with sensitive data, especially in cases where the effectiveness of anonymisation techniques was uncertain (Vadhan et al., 2012). The legal and policy tools designed included custom policies, contracts, licenses and other legal agreements tailored to the specific needs of researchers (and their data subjects) working with specific types of data under different technical approaches. The role of the legal instruments was to ensure the consistent application of computational tools in the appropriate cases and provide additional protection when necessary, such as transparent communication of data subject's consent, proper limitations on data use and sharing, audit mechanisms for regulating compliance etc.

A regulatory framework needs to be built for balancing the potential benefits of data processing operations against the potential harms of particular practices. Such a framework may include a set of best practices for data controllers articulating in non-technical terms the guarantees provided by the various data protection mechanisms and determining the parameters to be considered in evaluating whether certain processing activities can provide sufficient privacy without considerably affecting the data structure. It is essential that regulators compare the risks pertaining to processing operations with the benefits of unfettered information flow and solely in case the costs significantly outweigh the benefits of information flow, regulators should place restrictions on the storage, use and sharing of personal data. The ideal scenario would be to achieve maximum data protection and maximum data utility at the same time, but the realistic scenario is to achieve an acceptable trade-off, that is, the point at which the threshold for law should be established. The bottom line, however, is that meaningful discussions on addressing the utility-privacy trade-off require first bridging the gap between the different perspectives of technologists and lawmakers. To this end, it is crucial to achieve an understanding of the relation between the mathematical and the legal notions of privacy.

In an attempt to address the utility-privacy trade-off, the GPDR introduced a novel concept into the EU legal framework – the concept of 'pseudonymisation' – as a data security method aimed at protecting personal data while also allowing data controllers and processors to benefit from the data's utility (ICO, 2012a). 'Pseudonymisation' serves as an umbrella term

which encapsulates procedures such as data masking and hashing that are implemented in order to secure data directly or indirectly revealing an individual's identity (Art.29 WP, 2014a). Article 4(5) defines pseudonymisation as "the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information", while Recital 26 states that data which have undergone pseudonymisation "should be considered to be information on an identifiable natural person." As indicated, applying the pseudonymisation technique results in converting data about an identified individual into data about an identifiable person, which falls within the scope of the EU data protection law. The condition to pseudonymise a dataset is to keep the additional information necessary for re-identification safely inaccessible for the users of the data that have undergone pseudonymisation (Article 4(5)). The use of pseudonymisation techniques is supported by the GDPR as a privacy enhancing tool because it is recognised as being able to "reduce the risks to the data subjects concerned" (Recital 28). By reducing the linkability with the original identity of an individual pseudonymisation helps minimising the processed information and subsequently the risks presented by the processing operations. Pseudonymisation is used to disguise identities and enable the collection of data relating to the same individual without having to know their identity.

Notwithstanding the fact that pseudonymisation techniques - in the sense of removing or masking direct identifiers[81] - are considered to be an effective security method, depending upon context it is likely that they are not alone sufficient to adequately protect personal data. Rather they should be viewed as a part of an organisation's broader security policy encompassing a variety of security and control measures (Art.29 WP, 2014a; ICO, 2012a). The foregoing approach aligns with Recital 28, according to which pseudonymisation "is not intended to preclude any other measures of data protection", and Article 32(1) whose list of possible security measures to be applied by data controllers and processors is not limited to pseudonymisation. Not only can the use of pseudonymisation techniques assist data controllers in meeting their security obligations under the GDPR, but it may also offer the chance to ease the compliance burden. Despite the fact that data that has undergone pseudonymisation is still subject to data protection regulation, its legal effects appear beneficial to organisations handling personal data. Contrary to the DPD, where data that had

[81]Even though it seems that, within the meaning of Article 4(5), implementing pseudonymisation requires more than the mere removal or masking of direct identifiers, *See* Sophie Stalla-Bourdillon and Alison Knight, 2016, Anonymous data v. Personal data–A false debate: An EU perspective on anonymisation, pseudonymisation and personal data, *Wis. Int'l LJ.*.

undergone pseudonymisation was treated in the same manner as raw personal data and, as a consequence, companies lacked motivation to invest in the implementation of pseudonymisation mechanisms, the GDPR aims to encourage companies to adopt such security measures by promoting the value and significance of pseudonymisation in its provisions and by creating business incentives (Recital 29). Data controllers utilising pseudonymisation techniques can benefit from some level of flexibility as the GDPR relaxes several requirements on them. To begin with, the use of pseudonymisation allows organisations more leeway to process data beyond the original collection purposes as it is regarded a positive factor when the data controller determines whether further data use is "compatible" with the original use for which data were initially gathered (Article 6(4)(e)). Second, pseudonymisation constitutes a means of complying with the data protection by design obligation because it is a recognised process to demonstrate that adequate technical and organisational measures have been in place during the entire lifecycle of a system or product development (Article 25(1)). Third, data subjects may not need to be informed of a data breach affecting data that has undergone pseudonymisation in case the identification key that would allow re-identification has not been compromised (Article 34 (3)(a)). Fourth, the GDPR encourages data controllers to adopt codes of conduct that promote the use of pseudonymisation as a way to comply with the Regulation (Article 40(2)(d)). Finally, Article 89(1) suggests the application of pseudonymisation as an "appropriate safeguard" for the processing of personal data "for archiving purposes for the public interest, scientific or historical research purposes or statistical purposes.

The emphasis placed in the Regulation on the use of pseudonymisation means that organisations need to invest in the technologies required to prevent data from being in fully identifiable form; organisations handling personal data would be advised to pay serious attention to the possible uses of this technique throughout their business in order to minimise their liability under the Regulation. The pseudonymisation technique involves removing or obscuring direct identifiers – information that can be used to directly identify an individual such as name or social security number – so that the linkage to one's identity is rendered impossible without the use of additional information (Polonetsky et al., 2016). That is why, it is imperative that the additional information likely to lead to re-identification be held separately from the processed data and be subject to tight security controls to ensure non-attribution. In addition, data controllers should ideally prevent indirect identifiers– information able to reveal one's identity only if combined with other information such as

gender or data of birth - from being combined and used (Garfinkel, 2015; Rubinstein & Hartzog, 2016). This is because under Art. 4(5) GDPR it shall not be possible to attribute the data that has undergone pseudonymisation to an identified or identifiable individual. Although pseudonymisation can significantly reduce the risks associated with data processing, the fact that the data held in a separate database could be linked to the de-identified database through the use of an identification key raises the issue of the re-identification risk. To securely separate pseudonymous data from the identification key and thus mitigate the re-identification risk, data controllers should implement technical (e.g. tokenisation, hashing) and organisational measures (e.g. development of fully documented policies, personnel training) able to prevent the "unauthorised reversal of pseudonymisation" (Recital 75). In assessing whether a re-identification method is "reasonably likely to be used", organisations should have regard to "all the objective factors, such as the costs of and the amount of time required for identification (…), the available technology at the time of the processing and technological developments" (Recital 26).

Nevertheless, the integration of the pseudonymisation concept in the text of the GDPR appears to raise a range of issues that undermine the initial aspiration of the legislator to effectively address the utility-privacy trade-off. Even at the preparatory stage of the Regulation's drafting, the pseudonymisation concept was "a major area of divergent interpretation" as certain EU Member States (e.g., Austria, Germany, Ireland, Netherlands and UK) adopted the so-called relative approach considering encoded or pseudonymised data as identifiable, and hence personal data, "in relation to the actors that have means (the 'key') for identifying the data, but not in relation to other persons or entities", whereas other Member States (e.g., Denmark, France, Italy, Spain, Sweden) considered as personal any data that could possibly be linked to an individual by any third party, even in the hands of someone who had no reasonable means for such re-identification (absolute approach) (European Council, 2012). It is also noteworthy that only the Council's text formally recognised pseudonymisation as an example of appropriate security measures that data controllers and processors should apply to safeguard personal data (EDPS, 2015). Even though the introduction of the pseudonymisation concept signals a welcome shift from a binary to a tripartite approach to identifiability, and thus can be viewed as a positive development in incentivising organisations to use this technique as part of their overall compliance strategy by providing a safe harbour from certain data protection obligations, the answer to the question of when data has been pseudonymised seems to be somewhat

challenging. The conditions to be met are not clearly specified in the GDPR, which refers to the need of deploying measures able to prevent data from being attributed to an identified or identifiable individual without providing clear guidance as to the appropriate technical and organisational measures required to ensure that pseudonymised data has met the regulatory standards. The same holds true as regards the interpretation of how to treat pseudonymous data under the GDPR since the flexibility afforded to pseudonymous data is not either specified precisely. In addition, the fact that pseudonymisation is recognised as an example of security measures is likely to reduce the need for the implementation of other security measures based on the wrong assumption that the totality of measures taken to protect personal is enhanced, which does not correspond to the real effects of pseudonymisation given the high degree of re-identification risk involved in pseudonymisation techniques (Hintze, 2016). More importantly, the GDPR adopts the same static approach to pseudonymisation as with anonymisation, which ignores the fact that the re-identification risk of data processing operations is dependent on the interplay between the different components of data environments and should be continuously assessed (Elliot et. Al, 2016; Stalla-Bourdillon & Knight, 2016).

8.3.2. Need for a regulatory model establishing a data identifiability spectrum

Given that the concept of identifiability operates as a forceful legal trigger in the EU data protection legal framework, discussions regarding the scope of 'personal data', that is, whether a unique identifier meets the definition of 'personal data', are described as an all-or-nothing debate, which means that data is either subject to the full range of obligations under data protection law when relating to an identified or identifiable data subject, or it falls outside its scope and thus it is subject to none as is the case of anonymous data where data subjects are not or no longer identifiable. The fact that de-identification serves as "a tool that organisations can use to remove personal information from data that they collect, use, archive and share with other organisations" (Garfinkel, 2015) illustrates the significance of the de-identification process, that is, the process of modifying personal data to ensure that data subjects are no longer identifiable. De-identified data constitutes a vital aspect of digital economy as it enables organisations, governments and researchers to collect, store, use and share data for a broad range of purposes without jeopardising the privacy interests of data subjects. Furthermore, today governments and large organisations leverage the advances in Open data technologies and, as a result, they often publicly release large datasets in order to promote the public good. Notwithstanding the benefits of the de-identification process,

infamous re-identification cases demonstrating that de-identified datasets still remain vulnerable to re-identification attacks, some of which have been analysed above, have raised doubts about the underlying validity of de-identification and the extent to which it remains a credible method for using and deriving value from large datasets while protecting privacy. The consequence of those doubts has been the emergence of the so-called 'de-identification debate', which refers to the division of technical and legal experts on the efficacy of de-identification techniques (El Emam & Arbuckle, 2014). On the one hand, defenders of de-identification – the so-called "pragmatists" – argue that despite the theoretical and demonstrated ability to mount re-identification attacks, the re-identification risk in properly de-identified datasets remains minimal owing to the difficulties encountered in gaining access to auxiliary information, which means that even if data subjects may be distinguishable, they cannot be identified on an individual basis (Cavoukian & Castro, 2014). On the other hand, critics of de-identification – the so-called "formalists" - express the opinion that deploying de-identification techniques is impossible to eliminate privacy harms from publicly released data due to the increasing availability of background information that allows attackers to identify data subjects by mounting linkage attacks (Narayanan & Felten, 2014). The fact that discussions between pragmatists and formalists are often driven into definitional dead-ends is the legal implication of defining personal data as identifiable or not. The most worrying ramification of the de-identification debate, though, is the legal uncertainty to which it leads as it renders policymakers unable to judge whether current de-identification requirements should be maintained, reformed or abandoned. In turn, the lack of legal certainty around de-identification has broader consequences as it not only affects privacy-driven organisations, which structure their compliance strategies premised on the identifiable/non-identifiable distinction, but it also has implications on the data processing in research and data release of open data since legal uncertainty exacerbates the problems faced by researchers with respect to the steps that need to be taken to protect individuals' privacy before data processing.

To overcome the aforementioned problems stemming from the binary distinction between personal and non-personal data, which does not seem to accurately reflect the way data is treated in practice, the nature of the de-identification debate should be altered to a more nuanced and productive discussion on the obligations that should apply depending on the identifiability state of the data. Instead of abandoning the concept of identifiability, it would be rather better to abandon any strictly binary distinction between identifiability and non-identifiability, and consider both the concepts of identifiability and anonymity as end-points

on a wide spectrum with many interim points that pose a variety of graduated privacy risks (Rubinstein, 2016). Explicitly recognising that de-identification involves a broad spectrum of practices, each of which has different levels of strength, and identifying certain key points along that spectrum, which would establish a threshold, may have significant regulatory and policy implications as it enables the development of regulatory guidance that brings not only regulatory relief but also encourages the maximum use of de-identification compatible with the purposes of the data processing. This much-needed guidance tied to the levels of de-identification can, in turn, help achieve the optimal balance between maintaining utility of data and protecting individuals' personal data. If the processing of all types of personal data remains restricted to the same level, no matter whether additional protections have been imposed on certain datasets, then data custodians will be deprived of any incentives to invest in such protections; by contrast, simplifying, for example, the data release process will constitute a strong motivating factor in the broad implementation of such protections. On the condition that it becomes extremely difficult to link the data with the particular individual to whom the data relates, there is a strong argument that the risk of a privacy intrusive action impacting the individual is significantly reduced. When robust operational and contractual conditions have been established by companies in order to minimise the likelihood of re-identification, this should be acknowledged and incentivised. Otherwise, imposing disproportionate burdens on organisations processing data in a way that poses a limited risk to the protection of the data would be counterproductive for the law.

The different states of data would allow for broader uses of certain categories of data under certain conditions. Data that meets these conditions should be considered to have a significantly lower risk or re-identification and thus should be treated with greater flexibility in terms of its processing (e.g. being shared or sold for secondary purposes without having to obtain consent). As long as the re-identification risk falls below a specific threshold, data should still be regarded personal; however, in the cases when the risk is considerably reduced and is close to the 'de-identified' threshold, the law could justifiably add some flexibility to the processing of such data. When the risk is low, the law must be relaxed and allow for different types of consent or adapt collection, disclosure and use restrictions depending on the actual state of the data. It would be disproportionate to require an organisation, which has implemented the necessary measures, to comply with the full remit of data protection compliance obligations. Yet, regulators need to be careful as to the type of flexibility granted. For instance, data may be protected with respect to its processing and hence restrictions

placed on its use and disclosure could be raised, but this may not necessarily hold true in the event of a data breach where the risk to the rights and freedoms of the individuals affected may require the notification to them despite the safeguards applied to their data.

8.3.3. Existing literature on the degrees of data identifiability

The most crucial factor in assessing the risks presented by the processing of personal data is the level of de-identification of the data collected and processed because all risks are significantly reduced when de-identification is applied. De-identification involves a wide spectrum of data treatment techniques with different levels of strength. Each greater level of de-identification provides more data protection and further reduces the risk posed to individuals; therefore the stronger the level of de-identification, the lower the risk of re-identification. De-identification serves as a practical and useful tool for providing the much-needed clarity because it helps interpret the security requirements and resulting compliance obligations included in the data protection laws. Given that each state of identifiability poses a variety of graduated privacy risks, the regulatory and policy implications differ and hence the security measures that should be implemented vary depending on the level of de-identification. More specifically, the absence of any de-identification techniques should trigger more robust protection mechanisms, whereas relatively modest measures should suffice in case that strong de-identification is applied to data.

8.3.3.1. Hintze (2016)

Based on the answers to three questions (whether data is directly linked to identifying data, whether there is a systematic way of re-identification, whether data relates to a specific individual), Hintze describes the following four levels of identifiability as they are implicitly or explicitly described in the GDPR: *identified data, identifiable data, Article 11 de-identified data and anonymous or aggregate data*. To begin with, the GDPR definition of what constitutes 'personal data' provides the basis for the distinction between identified and identifiable data; in case the individual cannot be identified, identifiable data does represent a level of de-identification. Contrary to identified data, which is directly linked to an individual, identifiable data may relate to a certain individual whose identity is not apparent from the data. The GDPR explicitly recognises an additional level of de-identification by introducing the concept of pseudonymous data, that is, personal data that cannot be attributed to a specific individual without the use of additional information (which must be kept separately and subject to technical and organisational measures). Pseudonymous data is

considered to be a subset of identifiable data. Furthermore, the third level of de-identification is implicitly described in Article 11(1), which refers to data that shares the same qualities with the above category of identifiable data, insofar as an individual cannot be directly identified by the data, but the difference lies in the fact that this type of data could potentially be subject to re-identification if matched to additional information provided by the data subject. However, there is no systematic way for the data controller to reliably (re)create a link with the identifying data. This category of data is likely to enable the linking of some data to an identifiable individual with some degree of confidence; such were the datasets in the aforementioned re-identification cases of AOL and Netflix Prize, where data was erroneously characterised as anonymous and thus was publicly released. As set out in Articles 11(2) and 12(2), this level of de-identification determines the data controller's obligations under other provisions of the GDPR. Finally, the highest level of de-identification is reflected on the category of anonymous or aggregate data which, unlike the other three categories, does not relate to a specific person. Any foreseeable possibility of linking this data to an individual is eliminated because the data that could identify the individual has been removed or aggregated with other data in such a way that it does not contain any individual-level entries or events linkable to a specific person. Applying the rationale described above, the four levels of identifiability should trigger security obligations proportionate to the risk presented in such a way that the first two levels would call for stronger protections, the third level for some protections, while there would possibly be no security requirements for the last level of identifiability.

8.3.3.2. Polonetsky et al. (2016)

A more granular approach to the identifiability spectrum is suggested by Polotensky et al. who distinguish ten gradations of identifiability; data is divided into four broad categories – *personal data with different degrees of identifiability, pseudonymous data, de-identified data, anonymous data* – and each category is further divided into sub-categories. The primary variable for the classification of data is the treatment of *direct and indirect identifiers* which can be intact, partially masked, eliminated or transformed. As it becomes apparent by the terms, the difference between the two types of identifiers is whether data about an individual can identify the individual either directly or indirectly. In particular, *direct identifiers* are "data that can be used to identify a person without additional information or with cross-linking through other information that is in the public domain". Direct identifiers, such as name or social security number, can be protected by being suppressed or replaced with

symbols, generic names or random values. A direct identifier that is consistently replaced with a specific value becomes *a pseudonym* allowing linking personal information across multiple data records as long as similar direct identifiers are systematically pseudonymised. In contrast to direct identifiers, *indirect identifiers or else quasi-identifiers* enable a person's identification by helping connect various pieces of information about that individual; the combination of additional indirect identifiers results in one's identity emerging. Indirect identifiers, such as gender, date of birth, ZIP code, play a crucial role in the trade-off between data protection and data utility since they generally include important information whose removal from the dataset may damage the utility of the dataset. Some common methods of addressing the indirect identifiers include techniques such as suppressing or removal, generalising values as sets, swapping data between individual records, perturbing or adding noise. Besides the nature of identifiers, another element taken into consideration is the existence of safeguards and controls in place with respect to the collection, use and dissemination of the data. *Safeguards and controls* refer to legal, administrative and technical measures aiming to prevent employees, researchers or other third parties from re-identifying individuals. Such measures encompass both internal controls (e.g. security policies, personnel training, data segregation guidance, data deletion practices) and external controls (i.e. contractual agreements that address the use and sharing of data and contain auditing rights in order to ensure compliance).

As described below, the rationale of this approach is that different combinations of the above variables account for the different categories of data thus generating ten levels of identifiability. Based on the interplay of these variables, the first category includes the one end of the identifiability spectrum, that is, the explicitly personal data, where no identifiers have been obscured, and two more sub-categories which share the same characteristic that only direct identifiers have been partially masked; however, the difference between potentially identifiable data and not readily identifiable data lies in the existence of safeguards and controls in the latter case. A particularly interesting category in terms of the regulatory implications that it should entail is the category of key-coded data, where direct identifiers have been replaced by a key in order to avoid unintended or unwanted re-identification. Depending on whose hands the data is in, key-coded data could be viewed as either personal data (in the hands of the key holder) or pseudonymous data (in anyone else's hands). Even the strength of the key plays a role in assessing the degree of identifiability (e.g. encryption-based key allocation likely to be reversed as opposed to randomly mapped keys

that reduce the re-identification possibilities). The category of key-coded data is indicative of the need to sustain data utility as this sort of data is extensively used in a range of circumstances, such as scientific and historical research, where limited re-identification is necessary; in such cases strong controls are in place to ensure that a restricted set of players can hold the key. It is essential for the law not to treat all key-coded data in the same manner, but rather recognise the difference between key-coded data in the hands of the curator who also holds the key, and the same data in the hands of a third party which cannot easily "unlock" it. Otherwise, third parties such as researchers would have to sacrifice useful data and apply more cumbersome de-identification standards. A strict reading of the language used in both the DPD and the GDPR regarding the criteria for evaluating the identifiability of an individual, which refer to "all the means likely reasonably to be used either by the controller *or by any other person*", suggests that the legislator does not seem to have taken into consideration the aforementioned distinction as the likelihood of re-identification by a third party who does not have the key is determined based on the capabilities of the party who first coded the data.

Contrary to potentially identifiable or not readily identifiable data, where direct identifiers are only partially masked, pseudonymous data does not contain any direct identifiers that can be used to link data across contexts because these identifiers are either eliminated or transformed. In case pseudonymous data is protected with safeguards and controls over their release, then the data moves further down the identifiability spectrum to protected pseudonymous data. It is noteworthy that this definition of pseudonymous data is not necessarily in accordance with the definition of the GDPR as data that the GDPR denotes pseudonymous may under certain circumstances fall under the last categories of personal data following the classification of data made in this approach. Moving on to the next category, both direct and indirect identifiers have been either removed or manipulated in such a way that there is no linkage between the data and the real-world identities; again, depending on whether safeguards on publication and use controls are in place, data falls either into the de-identified or protected de-identified category. Given that the term de-identified data has been the focal point of heated arguments, it should be clarified that this model accommodates a risk-based approach to de-identification meaning that it focuses only on those attacks that an organisation considers as truly feasible rather than on every possible attack vector. Finally, as with the previous category, both direct and indirect identifiers have been removed in the case of the last category of data, that is, anonymous data. The difference, however, lies in the

mathematical and technical guarantees put in place in order to distort the data and hence prevent re-identification. On the other end of the identifiability spectrum lies the data that are so highly aggregated that there is no need for additional safeguards or controls.

8.3.3.3. Levin & Salido (2016)

The model proposed by Levin and Salido for assessing the risks presented by the processing of personal data recognises that the main two factors to be considered in the design of data protection measures are the following: (1) the desired data usefulness, which correlates to the de-identification techniques applied to the data; and (2) the anticipated data sharing scenarios, which encompass the technical, organisational and legal data protection measures used between the data source (the creator of the original dataset or the owner of individual data) and the data user (the recipient or the user of the data). The possible combinations of de-identification techniques and data sharing scenarios generate different levels of potential risk to data subjects. The point of intersection between the two axes (as presented in the model) denotes the extent to which personal data is protected from disclosure by the data user. The authors choose the above two factors to form a practical framework for the first phase of the design of data protection measures given that they are independent from the content of data and thus can be applied across different industries unlike other data-dependent factors, such as the value and the sensitivity of data, that need to be examined in a certain industry or application context. Their aim is to help narrow the range of de-identification techniques likely to be considered by practitioners. In an initial stage, a practitioner would need to identify the desired data utility and the possible data sharing scenarios before taking into account data specifics that would require more thorough technical knowledge. Based on the evaluation of the first-phase results and the inputs from different internal stakeholders regarding the intended usefulness of the data, the acceptable level of risk and therefore the appropriate de-identification techniques, a more detailed data-centric risk-based analysis would then take place for each use case.

Briefly presenting the proposed model, the data sharing scenarios examined range from the cases where data is restricted to an atomic legal entity (an organisation or an individual) to the cases where raw data is collected from individuals, depending on whether data is accessible only to authorised parties or to the general public and whether access to the data is provided under a signed agreement or not. As far as the de-identification techniques are concerned, they are ordered according to the degree to which the data retains its structure and

content after being de-identified; for instance, when no techniques have been applied, data retains its original form and usefulness. The first category of de-identification techniques ordered accordingly share the characteristic that retain their records format with increasing data loss. To begin with, pseudonymised data retains its original usefulness but linking the pseudonyms to the original data subjects is likely to occur with different degree of likelihood depending on which pseudonymisation technique has been applied (with or without controlled de-identification). Although the loss of data usefulness is relatively small in the following three types of de-identification techniques - masking of identifiers, masking of outliers and selective identifiers, generalisation of selective quasi-identifiers - the main difference lies in the fact that the ability to link between records relating to data is significantly reduced. The last three techniques of the first category – randomisation of selective quasi-identifiers, implementation of K-anonymity model for quasi-identifiers, creating synthetic data - preserve neither the data usefulness nor the data truthfulness as they alter the value of the attributes and thus greatly decrease the re-identification likelihood. Moreover, the second category of de-identification techniques includes techniques that provide statistical results, such as computation of statistics and implementation of differential privacy server model, which guarantee that the presence or absence of any particular data cannot be inferred from the de-identified dtaset. Finally, the principal feature of the last category is that the technique described, that is, implantation of differential privacy local model, provides so strong data protection guarantees that it can be used to generate both effectively anonymous microdata and accurate statistics.

8.3.3.4. El Emam et al. (2016)

In order to define precise criteria for evaluating the different types of data along the identifiability spectrum, the authors consider five interdependent characteristics of data sharing transactions and analyse the so-called “six states of data”; each state of data reflects the type of data exchange currently taking place. Although their analysis is the result of observations on the disclosure and use of health data, the authors claim that the states of data identified and the criteria suggested may also prove to be useful in other domains as well. The five factors taken into consideration when assessing the likelihood of re-identification are the following: whether data custodians are able to verify the identity of the data custodian; whether masking techniques have been applied to the datasets; whether de-identification techniques have been applied to the datasets; whether contractual controls have been put in place (e.g. whether the data recipient has signed an enforceable contract); and whether

security and privacy controls have been established. Based on the above criteria, the authors categorise the data in six states, of which the first three states refer to personal data given that the re-identification risk remains high. The first category involves the "raw" personal data which has not been modified at all or very little. The difference between the second and the third category – "vanilla" and "protected" pseudonymous data - is that the masking of the direct identifiers has not been followed by the application of additional contractual, security and privacy controls. Despite the fact, however, that such controls have been put in place in the third category and therefore the re-identification risk has considerably been reduced, protected pseudonymous data is still considered personal data since it is still below the de-identification threshold, which concerns the cases where de-identification techniques have been implemented, that is, where indirect identifiers have been either modified or removed. Such is the case of the fourth category - "non-public release of anonymised data" – where the data release requires strong safeguards and the prior authorisation of the data recipient who needs to have a legitimate interest to use the data. The identity of the data recipient also has to be verified in the fifth category – "quasi-public release of anonymised data" – and a terms-of-use agreement ought to be signed to prevent any re-identification attempt; however, anyone is in a position to request access to the files with individual-level information (or else, "microdata"). On the contrary, in the last category – "public release of anonymised data" – anyone can have access to microdata without even asking for that. In the last three categories, de-identification techniques perturbing the quasi-identifiers have been applied and, as a consequence, data belonging to these categories is qualified to be treated as non-personal data. Nonetheless, the level of perturbation may vary from low (e.g. a date of birth converted to month and year of birth) to high (e.g. a date of birth converted to a five-year interval). The level of perturbation is also associated with the usefulness of data; for instance, when the perturbation level is high (public data), the data quality is degraded, whereas the opposite occurs in the case of non-public data. Depending on the level of perturbation as well as the strength of the controls in place, the re-identification risk also varies.

By introducing a new point on the identifiability spectrum, the authors attempt to find a balance between preserving data utility and protecting personal data from re-identification attacks. Therefore, they propose a framework which would allow broader uses of pseudonymous data under certain conditions. The rationale behind this framework is that the current legislative framework fails to provide the necessary incentives to data custodians so that they invest resources in the protection of personal data. Under GDPR, pseudonymous

data can be used for secondary purposes by the same controller but cannot be shared with another party without consent (except limited cases). If, however, no express consent of the individual data subjects was required for the disclosure and use of pseudonymous data for secondary purposes, the data release process would be simplified and this could be a strong motivating factor for data custodians to implement the necessary security measures. The new category introduced by the authors is called “flexible pseudonymous data” in the sense that some added flexibility could be afforded in its processing as it is considered to have a significantly lower re-identification risk than other types of pseudonymous data as long as the following three conditions are met: first, the processing of data throughout its lifetime should be automated (no human activity involved) and, once the processing is completed, the data should be destroyed; second, the secure sharing of data requires that any analysis performed on the data does not produce results that would leak identifying information to the end-user; third, data sharing should not involve sensitive data.

8.4. Proposed Data Security Regulatory Model

The proposed regulatory model aims at providing a clear framework to data controllers with regard to the assets that should be protected and the level of security required covering all the aspects of the security of personal data processing. The model adopts a hybrid approach to the regulation of data controllers’ security obligations, which combines high-level security standards, or else the outcome to be achieved, with prescriptive guidance – examples of security measures that should be implemented in order to achieve the required outcome. By laying down high-level security objectives, which could reasonably be applied to a wide number of organisations handling personal data, and providing a baseline, that is, a minimum set of acceptable compliance strategies that must be in place, the regulatory model aspires to provide legal certainty without hampering innovation since no specific technological designs are mandated. To ensure a common understanding of the technical details, it is essential to develop an outcome-based approach determining a set of security objectives and security measures that data controllers should be aspiring to reach as a minimum, predicated on the aim of data protection principles. The security objectives that data controllers should be required to achieve in order to adequately protect the personal data they process are divided into the following categories: risk management; human resources security; security of premises and supplies; security of networks; security of servers; security of communications;

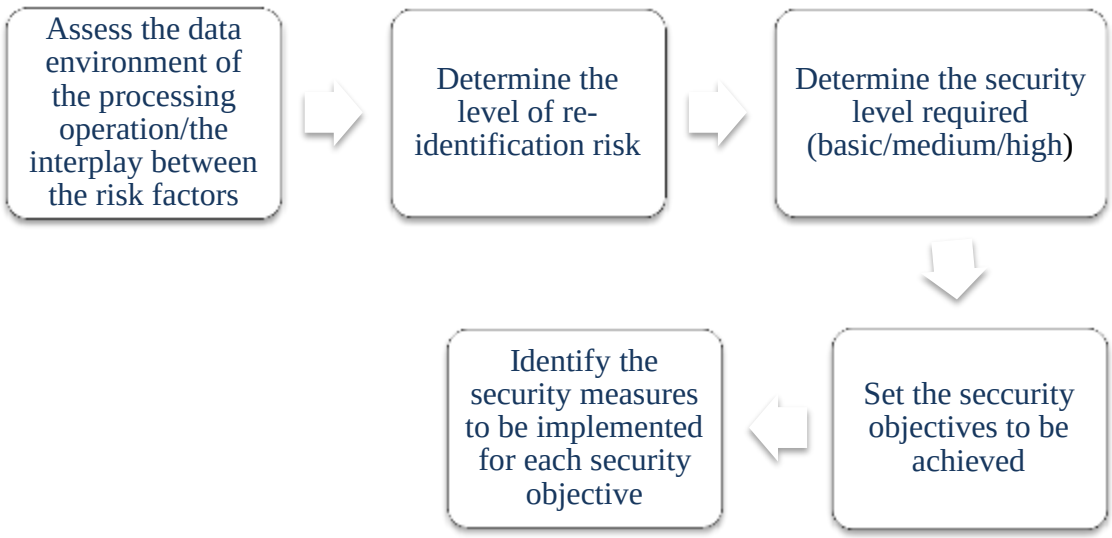
business continuity management; incident management; ongoing compliance monitoring. In assessing the appropriate level of the security measures that must be applied for each security objective, the risks presented by the processing of the data are taken into account.

The proposed regulatory model establishes a layered approach, according to which data controllers should apply different levels of security – basic, medium, high - depending on certain risk factors. A layered approach to data security suggests that regulators establish a sufficiently comprehensive model appropriate for different contexts that would be regularly updated with the input of experts and stakeholders. In particular, each entry in the model would specify a category of each one of the risk vectors. The aim of the model is its components to specify a set of contexts in which certain security requirements would apply and case-by-case review would be avoided in the initial phase of the design of security measures. More specifically, the application and interpretation of security requirements varies depending on the de-identification level as each state of data identifiability poses different privacy risks and creates different compliance obligations; the stronger the level of de-identification, the lower the risk posed to personal data. Accordingly, the security measures suggested rely on a continuum of identifiability choices, where data is completely identified, on one end, whereas re-identifying the data residing on the other end is extremely difficult. Therefore, emphasis needs to be placed on limiting access when data resides on one end, while controls over the dissemination of data should be put in place when data resides on the other end. In the between states of data, security policy should rather focus on the management of data.

Drawing on the literature on the data identifiability spectrum and the examples of the Spanish and Israeli data protection laws, the proposed model suggests a regulatory approach different to the one adopted in the current EU data protection regime. The dynamic approach suggested contrasts with the static approach of the EU legal framework, which distinguishes between personal and anonymous data ignoring the failure of the anonymisation technique to guarantee that data treated as anonymous in a certain context will remain anonymous under different circumstances. The approach suggested also differs from the static approach of the security legislation of Spain and Israel, which determine the security level required in each data processing operation according to predefined criteria that are merely associated with the nature of data, a certain number of persons with access credentials, or specific types of processing activities. Instead, this model introduces a dynamic approach to the required security level, which is contingent upon the data environment of each processing activity in

the sense that there is no fixed level of the re-identification risk involved in each processing but the risk is rather defined by the interplay between the risk factors. This model is different from the Spanish and Israeli models in one more aspect. Instead of stipulating technology specific security measures that may possibly favour certain technologies and become easily outdated, it lays down high-level security objectives and abstains from suggesting specific technological designs as examples of security measures. Furthermore, this model follows the same path as the authors that endorse the idea of incorporating various degrees of data identifiability into the data protection framework, but extends their work by systematically tailoring the security obligations of data controllers to the level of re-identification risk for different states of data and suggesting security measures proportionate to the re-identification risk level. The proposed model does not categorise the states of data according to specific factors, such as the data sharing scenario or de-identification techniques, but rather adds to the number of variables to be considered in assessing the re-identification risk and allows for the interplay between them to determine the state of data in each case. Two things should be borne in mind with respect to this model. First, this model solely serves as an example of how cyber security regulation should be formed. Second, it is essential that the model is regularly adapted to the technological changes not only in terms of the security objectives suggested but also with regard to the categories of data that should require special treatment. The same applies to the level of security required since what is currently considered to be a medium-level security measure may be a basic measure in ten years' time.

Figure: Steps involved in the proposed decision-making process



8.4.1. Risk factors

The list of factors presented is not exhaustive but rather indicative of the most crucial variables that need to be weighed by regulators in determining how protective data controllers should be when handling personal data. These factors can serve two purposes as they are both indicators of the risk level and instruments for reducing the re-identification risk. As indicators, they signal the likelihood and severity of potential privacy harm and hence enable both regulators and data controllers to appraise the level of the re-identification risk. The same factors can also be used to reduce the privacy risk by pointing to the level of security required in each case. By assessing the privacy implications of processing activities from the perspective of their impact on individuals, the likelihood of serious harm can be reduced. The results of such an assessment can be reflected in a better-targeted regulation providing for the necessary safeguards. In addition to the guidance provided in this way with respect to the security measures that must be implemented by data controllers, the use of risk vectors can also address the trade-off between data utility and data protection by providing regulators with flexibility in terms of the safeguards required in certain types of processing operations. For example, regulators might decline regulating private uses of data while placing restrictions to the public release of a certain category of data. Other factors that should be taken into account when assessing the risk of re-identification include, amongst others, the way the processing is structured, the interests at stake for the individuals, the advantages expected by the controller, the costs of conducting identification, possible technical failures, the risk of organisational dysfunctions such as breaches of confidentiality duties. As far as technological advances are concerned, one should consider the state of the art at the time of the processing as well as the possibilities for development during the period for which the data will be processed. The main risk factors that are considered to affect the likelihood of re-identification and therefore need to be balanced in determining the security measures a data controller must implement are the following: the value of data (e.g. the financial element increases the re-identification risk); the sensitivity of data (the possibility of redefining the concept of ‘sensitive data’ is examined in order to also include other forms of information such as sensitive data hidden in unstructured databases, some sorts of metadata, precise geolocation data etc.); the volume of data (e.g. large datasets entail high possibilities of re-identification attacks); the recipient of data (e.g. different security requirements should be imposed when data is released to trusted third parties or to the general public); the intended use of data (information to be processed for more problematic purposes require

stronger protection); the data treatment techniques (risk varies according to the way data is manipulated).

i. *Data Value:* The value of data is associated with the level of incentives an entity would have to use the data for purposes other than the approved and therefore correlates to the amount of resources an entity would be willing to invest to re-identify the data. In this regard, regulators should primarily weigh economic incentives for re-identification since financially-motivated attackers are more likely to comprise massive databases containing financial records and thus target a great number of individuals simultaneously.

ii. *Data Sensitivity:* The sensitivity of data is strongly connected to the amount of harm caused in the case of re-identification affecting both the data subjects (based on the content and the amount of detail included in the data) and the data source (depending on the number of the data subjects in the dataset). When there are files in an information system that require the application of a level of security measures different to that of the main system owing to the sensitive nature of the data they contain, they may be separated from the latter, with the relevant level of security measures being applicable in each case.

iii. *Data Volume:* The volume of data may affect the risk of re-identification as access to a large amount of information indicating personal preferences and behaviours enables attackers to match data to outside information and hence re-identification risk increases when data administrators store massive quantities of information. The fact that some large datasets have a high degree of unicity makes it easier to launch re-identification attacks as evidenced by re-identification cases where attackers were aided by the size of databases. Therefore it is important that law also regulates data quantity and not only data quality as regulators usually draw distinctions based on the types of data processed.

iv. *Data Use:* The intended use of data correlates to the motivation created for potential attackers to attempt re-identification. Some uses of data are less risky as is the case of data used for routine administrative purposes like customer service, website development or record keeping, and some other are more problematic

(e.g. when data is used for commercial or discriminatory purposes). Given the potential harm and motivations by attackers to identify individuals or sensitive attributes, information that is to be used for more problematic purposes must be better protected, whereas protections can be lowered when data is to be used to help people avoid serious harm or for a significant public good. Hence, regulatory requirements should be more relaxed in certain contexts, such as in the case when data is processed by a small number of actors who lack the motive to re-identify individuals in their datasets (e.g. fewer constraints can be imposed on researchers when processing data for purely research purposes). When there are files in an information system which require the application of a level of security measures different to that of the main system due to the their intended purpose, they may be separated from the latter, with the relevant level of security measures being applicable in each case.

v. *Data Treatment Techniques:* Re-identification risk is contingent upon the way data is manipulated through the use of de-identification techniques, which can, for example, suppress, generalise or hash data values to protect individual identities, or even create entirely new datasets that do not map to actual individuals and they are thus safer surrogates that authentic data. Even though the degree to which different data-handling techniques affect the re-identification risks cannot be defined with mathematical precision, computer scientists are able to roughly appraise whether that risk is high, medium or low.

vi. *Safeguards & Controls:* Re-identification risk varies according to the technical, administrative, organisational and legal controls placed on the use and release of information, which usually leverage the de-identification techniques to in order to more effectively safeguard the processing of data. For example, contractual agreements between initial data controllers and data recipients or a company's policies providing for meaningful and properly obtained consent can lower the re-identification risk and thus mitigate the need for robust protections. In terms of access controls, organisations may choose to release data only to internal staff or trusted recipients on the condition that they contractually agree to protect the data and refrain from attempting re-identification; recipient controls can be combined

with distribution controls; technical access control mechanisms can be utilised to limit who can have access to the data and how data can be accessed.

vii. *Data Recipient:* In broad terms, there are at least three different types of recipients of the data, each of which is considered to be increasingly risky – internal recipients, trusted recipients and the general public. In most cases, the first type of recipients is deemed less problematic, the second type riskier and the third type of recipients is seen as inherently problematic. The different level of re-identification risk involved in each type of recipient means that different protections are required in each case depending on the 'trust' criterion. For example, regulators can limit the flow of information to trusted relationships between private parties and scrutinise public releases of data to the general public much more closely as they require the greatest amount of protections. In addition, regulators can differentiate the amount of protections required even between trusted parties by introducing several tiers of trusted recipients, where data sharing with recipients at the lowest tier would be treated as the equivalent of public release while increasing protections would be tied to less trustworthy recipients. The 'trust' element also plays a crucial role in the trade-off between data utility and data protection as it encourages the use of either the authentic data or slightly modified data by trusted parties. For example, trusted third parties at research universities, which are less likely to use data inappropriately or attempt re-identification might warrant access to richer, less anonymised data for research purposes.

8.4.2. Reconsidering the category of sensitive data as defined in the GDPR

Focusing on the threshold definition of sensitive personal data, it appears that information is classified as sensitive not merely because it might conceivably lead to harm but because the probability of harm is regarded sufficiently high. Sensitive information describes information that if lost, compromised or disclosed could result in substantial harm, embarrassment, inconvenience, or unfairness to an individual. The risk of severe harm resulting from a loss of control over information justifies the fact that the law requires even more diligence for sensitive information commanding the custodian of such information to comply with special handling rules (of collection, use and disclosure) in order to prevent privacy or security harm. In determining whether certain information may trigger a risk of harm for the individual and hence considering whether personal information needs to be treated as sensitive data, the

criteria outlined in the work of Gratton (2013) will be briefly presented. To begin with, Gratton maintains that there are two types of information: information that may trigger a subjective type of harm and information that may trigger an objective type of harm. The first type of harm, which is associated with "injury to the feelings", would most likely take place when personal data is being collected or disclosed. Therefore, the disclosure of data is considered to be harmful in case it generates feelings of embarrassment or humiliation to the individual. However, it appears to be difficult to isolate the risk of harm as it is highly contextual upon disclosure; that is why, Gratton suggests the use of additional information that may be essential to the identification of the risk of harm. Such additional information that may help interpret the notion of "identifiable" in the light of the overall sensitivity of the information in question constitute the criteria for identifying the risk of harm. These criteria are the "intimate" nature of the information and the extent to which information is "available" to third parties or the public after the information is disclosed.

The criterion of 'intimate nature' includes, for example, information concerning health, information revealing political or religious beliefs, information that tend to reveal intimate details of the lifestyle and personal choices of the individual. Moreover, geolocation, metadata and criminal records may also be considered as information of intimate nature. For instance, the processing of personal data relating to criminal convictions and offences is subject to restrictions by Article 10 of the GDPR. As far as the 'availability' criterion is concerned, the degree of data sensitivity decreases once a given set of data is in circulation or available to the party receiving the information since the risk of subjective harm is less severe in this case; the disclosure of information about an individual that is already available is less likely to make the individual feel embarrassed or humiliated. Especially in the information age, where most information disclosed has already been available to a certain context, it is important to take into account the degree to which information is made more accessible in assessing its sensitivity. In terms of the information triggering an objective harm, this category includes information which, upon being used, may either directly cause harm to an individual or may lead to a negative impact for the individual behind the information. In this regard, objective harm would be triggered by the unanticipated or forced use of personal information against a given person as are the cases of identity theft, financial damages as a consequence of loss of business opportunities, damage or loss of property, negative effects on the credit record, physical harm or information inequality. Often, information is used by organisations for various purposes which may have no impact or a limited or indirect impact

to the individuals. Gratton articulates the view that in such cases information should less likely be considered as sensitive.

After engaging in a risk assessment taking into consideration the likelihood that certain types of information will lead to particular types of harm, the EU legislator describes as sensitive data the "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, data concerning health or sex life" (Article 8 DPD), while the GDPR also proscribes the processing of genetic data as well as the processing of biometric data "for uniquely identifying a natural person" (Article 9). However, even the most recent Regulation does not seem to cover all the possible categories of data that need to be treated as sensitive, which is indicative of the slow development of sensitive information law that often fails to keep pace with the technological advances. When accurately calculating the probability of harm, sensitive information seems to arise with alarming frequency. This frequency dictates the need to constantly develop new approaches to assessing the probability of harm to include new types of information whose nature requires special protection. To this end, it has been suggested that the law extends to also cover information whose processing could lead to significant harm given the technologies currently available, such as precise geolocation information, remote biometric data, metadata, and sensitive information stored in unstructured databases when such information can be easily extracted.

8.4.2.1. Precise geolocation data

The term 'location data' is likely to encompass any form of data that has a geographic position associated to it, such as information collected by smartphone devices, swipe cards, wireless networks for navigation, location-based advertising, or whereabouts tracking purposes. The value of location data lies in its ability to infer even more useful information about an individual than the face value of the original data (e.g. travel patterns, religion, health issues). According to the Interactive Advertising Bureau, "location data not only improves publisher's own advertising insights but also opens up completely new avenues of monetisation" which explains the great amounts of money spent on location-targeted mobile advertising. The feature that adds to the value of location data is also the one which renders its processing dangerous for individuals' privacy. An example illustrative of the privacy concerns linked to the processing of location data is the use of Mac addresses which can potentially result in the identification of an individual, especially in light of the significant

growth of the Internet of Things market, where Mac address can relate to a personal device such as a smartphone or a fitness tracker (Sumroy & Cousin, 2016). The last few years, there has been an explosion in the technology of precise geolocation tracking facilitating not only the collection of a larger amount of information about a greater number of individuals, but also the storage of the information for a longer period of time, its distribution to a broader group of people, and its processing for more purposes than in the past. The reasons behind such practices are primarily technological and economic. First, cheap GPS sensors embedded in small mobile devices, in combination with smartphone operating systems that disclose location information to application developers through an application programming interface (often location is made known without the consent of the user), result in mounting concerns regarding the potential privacy and security harms. Second, stronger economic incentives are provided as the wireless providers have commenced to monetise user location information (Ohm, 2015). This problem takes worrying dimensions when considering those having access to databases containing sensitive information, such as service providers, contractors, employees etc. What is more, the likelihood and the magnitude of a potential harm become even greater with the rise of the cloud services which provide access to a vast amount of information to the biggest online services whose data protection policies do not always contain the necessary safeguards to adequately protect personal data and whose systems are more susceptible to cyber attacks. Therefore, the case of precise geolocation data seems to meet the criteria for falling under the category of sensitive data given the severity of harm that is likely to be caused if placed in the wrong hands (Escudero, 2001). Even though the GDPR has taken a significant step to include location data within the definition of personal data (Article 4(1)), unlike the DPD which did not apply to this category of data, yet location data, and especially precise location data, is not treated as potentially (under certain circumstances) sensitive information under the EU law. At this point, it is worth mentioning the Spanish law implementing the DPD, which interestingly provides for the application of high-level security measures to traffic and location data processed by providers of public electronic communications networks or services (Article 81 (4)).

8.4.2.2. Remote biometric data

The spread of biometric data, such as fingerprints, facial recognition data, and iris scan, gives rise to two types of privacy harm. The first type of harm raises similar issues with those posed in the discussion of precise geolocation information as it refers to the ability of biometric techniques that work from a distance (e.g. facial recognition of images captured

through closed-circuit television) to record the precise location of individuals. The second type of harm materialises when biometric information is used to spoof identity (Ohm, 2015). In contrast to the past, when the idea of securing authentication to data or information by applying biometric techniques had not even been conceived biometric data such as fingerprints or face prints are widely used as more seamless and simpler means of access control. Given that biometric data is as sensitive as the information it protects, the more organisations embrace biometric access to protect valuable information, the more harm is caused by unauthorised access and, as a consequence, the more sensitive remote biometric data becomes. The processing of biometric data is explicitly protected by the GDPR as it falls under the term of 'personal data' but it is not included in those categories of data enjoying special treatment. It would be prudent for lawmakers to weigh not only current implementations of biometric data by system designers but also plans for future implementations in determining whether this category of data could potentially be treated as sensitive data taking into consideration the circumstances under which it is likely to be processed.

8.4.2.3. Metadata

For at least a decade, scholars have argued that the treatment of the metadata associated with communication technologies should not be differentiated from that of the content of communications on the grounds that metadata is simply 'data about data' and hence does not deserve full protection. As a matter of fact, not only can some sorts of metadata directly identify individuals but they can also reveal sensitive information about them. Although certain types of metadata solely contain the information necessary for the traffic flow (e.g. source and destination IP, source and destination port, protocol, sequence number), other types of metadata (e.g. email subject line, email address sender/receiver, domain name of a website) can potentially reveal more information about the individuals than the content of communications itself or help make inferences with respect to information that currently fall under the category of sensitive data under the GDPR.[82] The demand for strengthening the legal framework surrounding the protection of metadata attracted more attention after

[82] *See* Sophie Stalla-Bourdillon, Evangelia Papadaki, and Tim Chown, Metadata, Traffic Data, Communications Data, Service Use Information... What is the Difference? Does the Difference Matter? An Interdisciplinary View from the UK (May 2015), Serge Gutwirth & Ronald Leenes, Data Protection on the Move, Springer 2015, available at SSRN: https://ssrn.com/abstract=2625181.

Snowden's revelations which demonstrated the great amount of information included in metadata and how such information can be processed by government agencies for surveillance purposes. Considering the harm likely to be caused by the processing of certain categories of metadata, it appears that metadata also meets the threshold definition of sensitive data which places emphasis on the sufficiently high possibility of significant harm. The EU legislator seems to have recognised the crucial importance of metadata as evidenced by the reform process of the e-Privacy Directive. According to the proposal for a Privacy and Electronic Communications Regulation, electronic communications data "may also expose very sensitive and personal information, allowing precise conclusions to be drawn regarding the private lives of the persons, such as their social relationships, their habits and activities of everyday life, their interests, tastes etc." (Recital 2).

8.4.2.4. Unstructured data

Today, large amounts of sensitive information end up in giant, unstructured pools of information maintained by tech industry giants. For example, Google stores each search query in what has been named "the database of intentions" (Battelle, 2006), Facebook maintains a sensitive database of past photos and status updates, while email providers store repositories of sensitive messages (Grimmelmann, 2009). The common features of the practices applied by online companies lie in the fact that the information is collected incidentally, not in a targeted manner, and that they are blended with less sensitive and non-sensitive information. In the recent past, database owners who possessed vast stores of sensitive information could render them relatively inaccessible in unstructured data. Limits in computation, storage and programming tools and techniques made it impossible or at least particularly difficult for anyone to manage to identify an individual in such databases since sensitive information could easily hide in the crowd (Ohm, 2009). However, the shift in technological possibility brought about by the recently developed and still emerging data analytics has enabled technologists to parse out sensitive information and find ways of imparting meaning and structure to unstructured data. Big data and the associated analytics tools, coupled with the emergence of cloud, mobile and social computing offer opportunities that enable greater understanding of complex infrastructures. For instance, it has been proved that Google can convert billions of individual searches in the form of unstructured data into meaningful data, which poses the question of whether search queries should be treated as unregulated unstructured data or fully recognised sensitive information (Ginsberg et al., 2009). Illustrative of this concern is the example of Google's massive database of search

queries containing health information and yet the database itself is not considered to be sensitive health information. Current data protection laws tend to focus on the type of information deemed to be sensitive, not the structure of the database in which the data is found. As a matter of fact, unstructured data fall outside the scope of the GDPR which applies only to the processing of personal data that "form part of a filing system or are intended to form part of a filing sytem" (Article 2 (1)); as filing system is defined "any *structured* set of personal data which are accessible according to specific criteria, whether centralised, decentralised or dispersed on a functional or geographical basis" (Article 4 (6)). As stated by Solove (2008), it is imperative that the law recognises that the digitisation and searchability of records that previously existed only as paper copies, or else "the problem of obscurity" (Hartzog & Stutzman, 2013), can lead to new privacy harms. Although it would not be wise to create a law that treats all search queries or email messages as sensitive data, the law should be expanded to cover the cases, such as the abovementioned case of Google, where the surrounding context suggests that a company has both the tools and incentives to parse out the sensitive information hidden within larger pools of data, commingled with non-sensitive information. In such cases, not only should data such as search queries be treated as regulated unstructured data but also as fully realised sensitive information (e.g. when containing health information).

8.4.3. Examples of the interplay between the risk factors and the re-identification risk level

Table 1: Examples of processing activities involving low re-identification risk

Risk Factors	Low Risk Level	Low Risk Level	Low Risk Level
Data Value	Data with no economic value	Data with little economic value	Data containing financial records
Data Sensitivity	Data not pertaining to any of the categories of sensitive data	Location data or metadata	Data pertaining to any of the categories of sensitive data
Data Volume	Databases containing information about a large number of data subjects	Databases containing information about a small number of data subjects	Databases containing information about a small number of data subjects
Data Use	Data processed for routine administrative purposes	Data processed for navigation purposes	Data processed for non-commercial purposes
Data Treatment Techniques	Both direct and indirect identifiers have been eliminated or transformed	Both direct and indirect identifiers have been eliminated or transformed	Both direct and indirect identifiers have been eliminated or transformed
Safeguards and Controls placed on data	Access controls and legally binding contracts in place	Access controls and legally binding contracts in place	Data must not be used to make information available to third parties
Data Sharing	Data shared to the general public	Data shared among trusted parties	Data shared among internal recipients

Table 2: Examples of processing activities involving medium re-identification risk

Risk Factors	Medium Risk Level	Medium Risk Level	Medium Risk Level
Data Value	Data with no economic value	Data with little economic value	Data containing financial records
Data Sensitivity	Data relating to criminal offences	Data regarding an individual's habits that may be indicative of any of the special categories of data.	Data pertaining to any of the categories of sensitive data
Data Volume	Databases containing information about a large number of data subjects	Databases containing information about a large number of data subjects	Databases containing information about a small number of data subjects
Data Use	Data processed for administration purposes	Data processed for commercial purposes	Data used for the provision of financial purposes or tax administration purposes
Data Treatment Techniques	Both direct and indirect identifiers have been partially masked	Both direct and indirect identifiers have been eliminated or transformed	Both direct and indirect identifiers have been partially masked
Safeguards and Controls placed on data	Access controls and policies on data processing in place	None or limited controls in place	None or limited controls in place
Data Sharing	Data shared to the general public	Data shared among trusted parties	Data shared among internal recipients

Table 3: Examples of processing activities involving high re-identification risk

Risk Factors	High Risk Level	High Risk Level	High Risk Level
Data Value	Data with no economic value	Data with little economic value	Data containing financial records
Data Sensitivity	Data relating to criminal offences	Data regarding an individual's habits that may be indicative of any of the special categories of data.	Data pertaining to any of the categories of sensitive data
Data Volume	Databases containing information about a large number of data subjects	Databases containing information about a great number of data subjects	Databases containing information about a large number of data subjects
Data Use	Data processed for administration purposes	Data used for automated profiling purposes	
Data Treatment Techniques	Direct identifiers have been partially masked whereas indirect identifiers remain intact	Both direct and indirect identifiers remain intact	Both direct and indirect identifiers remain intact
Safeguards and Controls placed on data	Access controls and policies on data processing in place	None or limited controls in place	None or limited controls in place
Data Sharing	Data shared to the general public	Data shared among trusted parties	Data shared among internal parties

8.4.4. Security objectives

Explaining high level security objectives provides the necessary guidance as to the outcome to be achieved for data controllers and processors, which becomes even more clear and structured when security is divided in different domains, whereas by determining a minimum set of security measures that must be in place in each domain, a baseline is also provided. Given the diversity in each business model, on which each organisation's security policy is based, the law should abstain from stipulating overly specific security requirements, which would fail to capture the particular features of each organisation involved. On the contrary, high level security objectives could be reasonably applied to a wide number of organisations. Such guiding principles represent a series of principles that all controllers and processors should be aspiring to reach as a minimum. In order to deal with the downside of laying down high level standards, which inevitably leaves a lot of significant technical details unaddressed, prescriptive guidance would be necessary to clarify the types of security measures that correspond to the risk level involved in every processing of personal data without, however, mandating specific technological designs. Below is presented a set of essential recommendations arranged in domains; per security objective, are listed security measures that could be implemented to achieve the objective described. Security measures are grouped in three levels of increasing sophistication.

1. Risk Management

Establishing and maintaining an appropriate governance and risk management framework requires performing risk assessments in order to identify and address the risks posed not only for the provider but also for the users who rely on the networks and communications services provided by the provider. Risk assessment allows an objective decision-making process as it offers a means of determining which assets are in scope and which security measures are appropriate for the risks presented. Data controllers shall perform risk assessments, specific for their particular setting, and regularly update the assessment to develop a cyber security policy tailored to the current risks and relevant threats.

2. Documented Security Policy

Establishing and maintaining an appropriate security policy requires that data controllers and, where applicable, data processors draft a document containing details relating to the scope of application of the document with specifications of the protected resources, the identity of the controller or the processor, the types of data collected, the estimated retention period of the

data, the intended usage of the data, the categories of recipients το whom the data will be disclosed, the categories of data subjects, the categories of personal data, the applicable regulations, codes and standards, the tasks and obligations of the personnel with access to critical premises, the procedure of data breach notification, a security incident response plan, the procedure for maintaining backup copies, a data recovery plan. The outcome of the risk assessment as well as the technical and organisational measures considered to be appropriate in terms of the risks identified shall also be included in the document, whose content shall be adapted to the changes likely to affect the security of personal data. Documented security policy allows to share a common view on how security should be managed and serves as a reference point for the security aspects that need to be covered.

3. Human Resources Security

Data controllers and processors shall establish and maintain a structure of the security roles and responsibilities of an organisation's personnel and ensure that the personnel understand the security policy that affects the performance of their functions as well as the disciplinary process in case of violation of this policy caused by personnel. Data controllers and processors must ensure that the personnel have sufficient security knowledge and are provided with regular security training. Background checks on personnel are recommended if required for their duties and responsibilities.

4. Security of Premises and Supplies

Data controllers and processors shall establish and maintain the physical and environmental security of the systems and processes supporting the provision of electronic communications networks and services. The security of supplies involves the security of electric power, fuel or cooling. In order to effectively protect the premises housing personal data, controllers and processors shall restrict the access of the staff members only to those resources required for the performance of their tasks and implement mechanisms to detect and mitigate possible intrusions to critical equipment. In case personal data is stored in portable devices or processed outside the premises of the controller or processor, the controller's or processor's prior authorisation shall be provided before any processing of the data.

5. Security of Servers and Applications

Data controllers and processors shall implement robust security measures in order to ensure the security of servers, which constitute the most critical equipment housing large amounts of personal data. It is essential that controllers and processors establish and maintain

mechanisms able to install operating systems critical updates without delay and guarantee the secure administration of databases. Therefore, it is recommended that controllers and processors refrain from operating servers that house databases used for other purposes, utilise personalised account identifiers to access databases, update applications when critical flaws have been identified and corrected, take precautions in the event of software installation or updates on operating systems to ensure business continuity and data availability, use vulnerability detection tools in case software is being executed in servers.

6. Security of Network and Information Systems

Data controllers and processors shall establish and maintain logical access controls to guarantee the security of their internal network and also set forth contractual requirements with third parties to ensure that dependencies on third parties do not negatively affect the security of networks and services. Data controllers and processors must be able to ensure that each user can only access the data she needs for performing her mission. To that effect, access controls mechanisms shall be put in place consisting of authentication mechanisms, which will provide a unique identifier to each user, and authorisation mechanisms, which will apply prior access controls to data for each category of users. When the authentication mechanisms are based on the existence of passwords, there shall be a procedure of disclosure, distribution and storage guaranteeing their confidentiality and integrity, and the frequency with which the passwords shall be changed, which under no circumstances must be less than yearly, shall be defined in the security document. As far as the integrity of network and information systems is concerned, there shall be deployed security mechanisms able to provide defence in depth by protecting from viruses, code injections and other malware that can alter the functionality of systems. Data controllers and processors shall make sure that software of network and information systems is not tampered with or altered, security critical data such as passwords and private keys are not disclosed or tampered with, and frequent controls shall be conducted to detect malware on internal network and information systems. Data controllers and processors must evaluate and review the effectiveness of access control policies and procedures as well as the effectiveness of the measures deployed to protect the integrity of network and information systems.

7. Security of Communications

Data controllers and processors shall establish and maintain security mechanisms able to guarantee the confidentiality, integrity and authenticity of communications. Security measures should be put in place to safeguard personal data stored in databases and,

especially, data transmitted over the Internet taking into account the significant risk of the disclosure of such data. Such measures include, among others, the encryption and pseudonymisation of personal data as well as the use of enhanced security protocols able to ensure the authentication of servers. Data controllers and processors shall only collect the data necessary with respect to the purposes for which they are processed (data minimisation principle) and erase or destroy data no longer necessary, such as data stored in equipment to be discarded. The systems housing personal data must include a mechanism for suppressing, archiving or anonymising data when its retention period expires.

8. Business Continuity Management

Data controllers and processors shall establish and maintain contingency plans to ensure the continuity of the networks and communications services as well as recovery mechanisms for restoring network and communications services in the event of a physical or technical incident. To that effect, protocols for making backup copies of personal data must be established and copies should be tested on a regular basis. Procedures for the recovery of data shall be established to guarantee at all times the reconstruction of data to the original state at the moment the loss or destruction occurred. Data controllers and processors shall perform the backup of software used for data processing in order to guarantee its continued existence. Physical protection measures against natural disasters should also be considered.

9. Incident Management

Data controllers and processors shall establish and maintain mechanisms able to minimise the impact of security incidents on users. To that effect, there shall be established response procedures for timely detecting, assessing, managing and reporting incidents. Data controllers and processors must be prepared to act promptly to restrict access, isolate impacted information and systems, notify key partners and conduct rapid damage control. In order to effectively address a security incident, it is necessary to establish a register for recording the actions performed on the information processing systems, which shall be made available to the supervisory authority on request. The record shall contain information relating to the type of incident, the time of occurrence, or if appropriate, detection, the person making the notification, the person to whom the incident was communicated, the estimated impact of the incident, and the security measures applied to mitigate the impact. Data controllers and processors must guarantee that records cannot be altered, and in any case save the information contained in the records for a period of time that is not excessive.

10. Ongoing Monitoring

Data controllers and processors shall establish and maintain processes for regularly monitoring, testing and auditing network and information systems and facilities, and evaluating the effectiveness of security measures applied for ensuring the security of processing. To this end, controllers and processors must implement documented policies and procedures for monitoring including minimum monitoring and logging requirements, retention period, tools for monitoring systems and collecting logs, and the overall objectives of storing monitoring data and logs. More specifically, policies and procedures should be implemented to monitor compliance to standards and legal requirements, containing assets, processes and infrastructure to be audited, templates for audit reports, guidelines on the person in charge of carrying out the audit (internal or external) as well as the objectives of auditing. Policies and procedures for compliance and auditing should be reviewed and updated, taking into account changes and past incidents. As far as the assessment of technical and organisational measures is concerned, data controllers and processors can demonstrate evidence of compliance by adhering to approved certification mechanisms and data protection seals and marks or approved codes of conduct.

8.4.5. Examples of security measures per security objective

1. Risk Management (SO1)

Security Level	Proposed Security Measures
Low Security Level	▪ Risk assessment shall cover all aspects (both technical and non-technical) of an organisation. ▪ Risk assessment shall follow any methodology that identifies assets, threats and counter-measures. ▪ Risk assessment shall be reviewed in the following cases: at least every three years; when any significant internal change occurs; at least in six months after any external change occurs. ▪ The threats and risk of each processing operation shall be studied and prioritised according to their gravity and probability of occurrence.
Medium Security Level	▪ Risk assessment shall cover all aspects (both technical and non-technical) of an organisation. ▪ Risk assessment shall follow the methodology recommended by the national supervisory body. ▪ Risk assessment shall be formally validated by Management. ▪ Risk assessment shall be reviewed in the following cases: at least once a year; before any significant internal change occurs; at least in four months after a significant external change occurs. ▪ Genuine methods for studying risks shall be used to enhance the exhaustiveness and depth of risk study.
High Security Level	▪ Risk assessment shall cover all aspects (both technical and non-technical) of an organisation. ▪ Risk assessment shall follow the methodology recommended by the national supervisory body. ▪ Risk assessment shall be formally validated by Management. ▪ Risk assessment shall be reviewed in the following cases: at least once a year; before any significant internal change occurs; at least in two months after a significant external change occurs. ▪ A security audit of the information systems shall be conducted.

2. Documented Security Policy (SO2)

Security Level	Proposed Security Measures
Low Security Level	▪ Information relating to access privileges, authentication mechanisms, communication security and security incidents shall be documented and retained for two years. ▪ Each employee shall have access to written or electronic documentation for each operational procedure they have to perform as well as to the documented general policies.
Medium Security Level	▪ Each employee shall be personally notified of any change in the documentation.
High Security Level	▪ Documentation shall contain detailed information on relating to the data collected, the purposes of the processing activities, the types of data processed, the identity of the database manager, its data security officer and its holders (i.e. those who usually possess a copy of the database and are entitled to use it), as well as the processing activities carried out by other processors, the security risks involved and possible ways of mitigating them. ▪ Each employee shall be personally notified of any change in the documentation.

3. Human Resources (SO3)

Security Level	Proposed Security Measures
Low Security Level	▪ Each employee shall follow a security awareness session every three years. ▪ The content of the security awareness actions, including information, training and awareness of disciplinary sanctions, shall be reviewed after every major change in the security policy and, in any case, at least every three years. ▪ Employees with access credentials shall be provided with training with respect to the security protocols and security requirements before being granted access privileges.
Medium Security Level	▪ Each employee shall follow a security awareness session every two years. ▪ The content of the security awareness actions, including information, training and awareness of disciplinary sanctions, shall be reviewed after every major change in the security policy and, in any case, at least every two years.
High Security Level	▪ Each employee shall follow a security awareness session every year. ▪ The content of the security awareness actions, including

	information, training and awareness of disciplinary sanctions, shall be reviewed after every major change in the security policy and, in any case, at least every year.

4. Security of Premises and Supplies (SO4)

Security Level	Proposed Security Measures
Low Security Level	▪ Access to the premises shall be based on a single-factor and a formal identification.
Medium Security Level	▪ Access to the premises shall be based on two-factor authentication. ▪ Highly critical components shall be accessed with dual control. ▪ Access to databases' premises and equipment brought in or taken out of such premises shall be monitored.
High Security Level	▪ Premises where computers used for data treatment are located shall be subject to strict physical boundaries and control checks. ▪ Access to the premises shall be based on two-factor authentication. ▪ Highly critical components shall be accessed with dual control. ▪ Identification of the personnel accessing the premises shall be performed.

5. Security of Servers and Applications (SO5)

Security Level	Proposed Security Measures
Low Security Level	▪ Security requirements shall be taken into account as soon as the application is being designed. ▪ Applications shall be updated when critical flaws have been identified and corrected. ▪ Servers housing databases used for other purposes shall not be used for the storage of personal data.
Medium Security Level	▪ Updated security protocols shall be used during data transmission to ensure servers authentication. ▪ The data processing development of applications shall be carried out in a data processing environment separate from that of production. ▪ Sensitive systems, i.e. systems processing personal data or data considered as business confidential, shall have a

	dedicated (isolated) data processing environment.
High Security Level	▪ The number of applications requiring administrator level rights for their execution shall be limited. ▪ Vulnerability and attack detection tools shall be used when software is executed on servers.

6. Security of Network and Information Systems (SO6)

Security Level	Proposed Security Measures
Low Security Level	▪ Authentication shall be performed with software certificate or strong password otherwise. ▪ In case authentication mechanisms are based on passwords, stored passwords shall be encrypted and changed at least once a year. ▪ Information systems shall be subjected to penetration tests once in eighteen months in order to evaluate their robustness in the face of internal and external security risks.
Medium Security Level	▪ Authentication shall be performed with certificates stored within secured device (e.g. with a qualified certificate) or with software certificate otherwise. ▪ Automated mechanisms shall be established to monitor access to databases and logs shall be maintained for at least two years. ▪ Information systems shall be subjected to penetration tests once in twelve months in order to evaluate their robustness in the face of internal and external security risks.
High Security Level	▪ Authentication shall be performed with certificates stored within secured device (e.g. with a qualified certificate) each time it is possible, with software certificate or strong password otherwise. ▪ In case authentication mechanisms are based on passwords, stored passwords shall be encrypted and changed at least once a year. ▪ Data access control mechanisms shall be integral part of the data processing software development. ▪ Information systems shall be subjected to penetration tests once in six months in order to evaluate their robustness in the face of internal and external security risks. ▪ Networks shall be segmented into logical sub-networks according to the services deployed on them. ▪ Intrusion detection systems shall be installed to analyse real-time network traffic and detect suspicious activities.

7. Security of Communications (SO7)

Security Level	Proposed Security Measures
Low Security Level	▪ All directly identifying information shall be (wholly or partially) masked or removed from the dataset.
Medium Security Level	▪ The distribution of files containing personal data or the transfer of personal data through public or wireless electronic communications network shall guarantee that data unintelligible either by encoding them or using any other mechanism for this purpose. ▪ Both directly and indirectly identifying information shall be removed from the dataset.
High Security Level	▪ Transmitted data shall be protected by means of cryptographic software programs that have been verified by third parties with proven experience. ▪ In case data transmission is carried out by sending a physical media, cryptographic mechanisms shall be applied to the data before it is stored on media. ▪ Updated security protocols guaranteeing data confidentiality shall be applied to file transfers. ▪ Transmission of the secret guaranteeing the confidentiality of the transfer (cryptographic key, password etc.) shall be carried out through a separate transmission, if possible using a different type of channel from the one used for the transmission of the data. ▪ Both directly and indirectly identifying information shall be eliminated or transformed.

8. Business Continuity Management (SO8)

Security Level	Proposed Security Measures
Low Security Level	▪ Business continuity and recovery plans shall be tested at least every three years. ▪ Integrity of backup shall be checked at least once every month.
Medium Security Level	▪ Protocols for making backup copies and procedures for the recovery of data shall be established. ▪ The implementation of backup protocols and recovery procedures shall be monitored every six months. ▪ Business continuity and recovery plans shall be tested at least every two years. ▪ Integrity of backup shall be checked at least once every week.

High Security Level	▪ Business continuity and recovery plans shall be tested at least every year. ▪ Integrity of backup shall be checked after each backup.

9. Incident Management (SO9)

Security Level	Proposed Security Measures
Low Security Level	▪ Audit logs shall be monitored or reviewed at least once a month to identify evidence of malicious activity. ▪ Security incidents shall be reviewed at least once every nine months to determine the need for updating security protocols.
Medium Security Level	▪ Audit logs shall be monitored or reviewed at least once every week to identify evidence of malicious activity. ▪ Security incidents shall be reviewed at least once every six months to determine the need for updating security protocols.
High Security Level	▪ Audit logs shall be monitored or reviewed at least once every week to identify evidence of malicious activity. ▪ Security incidents shall be reviewed at least once every three months to assess the need for updating security protocols.

10. Ongoing Monitoring (SO10)

Security Level	Proposed Security Measures
Low Security Level	▪ An audit (internal or external) shall be conducted at least every three years to determine the need for updating documentation of security policy and security protocols.
Medium Security Level	▪ An audit (internal or external) shall be conducted at least every two years to determine the need for updating documentation of security policy and security protocols. ▪ A data protection officer in charge of monitoring the implementation of security measures and analysing the audit reports shall be appointed.
High Security Level	▪ An audit (internal or external) shall be conducted at least once a year to determine the need for updating documentation of security policy and security protocols. ▪ A data security officer shall be appointed to be in charge of establishing data security protocols and preparing ongoing plans to review compliance.

8.5. Conclusion

As the above analysis has demonstrated, the failure of the anonymisation techniques to sufficiently safeguard individuals' privacy signals the need for a new data protection paradigm which combines a risk-based approach to security obligations with different states of data on the identifiability spectrum. A regulatory model based on risks and outcomes enables organisations handling personal data to go beyond mere compliance with regulatory requirements and develop a better-structured approach to the processing of the data they collect, use and share. The risk of re-identification, coupled with the effects of data processing on the data subjects, can serve as points of reference for regulators to set a threshold, below which personal data should be considered as personal data under certain circumstances and hence treated accordingly. The state of data should be constantly re-assessed given the fact that current technologies can easily convert data considered to be de-identified into identified data. In addition, the point at which this regulatory threshold is established needs to be based on the realistic scenario that a certain level of risk is acceptable, which, in turn, requires a common understanding of the notion of the re-identification risk between technologists and lawmakers who often support the ideal, albeit unrealistic, scenario of a zero-risk approach to data protection. Therefore, in cases where the risk is considerably reduced and is close to the de-identified threshold, the law should add some flexibility to the processing of data in order to incentivise companies in charge of data processing to implement the necessary security protections. To achieve an optimal balance between maintaining the utility of data and protecting individuals' personal data, it is essential that the law recognises the existence of different degrees of data identifiability which pose threats of different significance, and tailor security obligations to the re-identification risk level involved in each processing operation. To this end, the regulatory model proposed can serve as an example of the way laws regulating the security obligations of data controllers should be drafted as it aspires to accomplish two goals – address the issues arising from the technology neutral approach to security obligations by providing clear guidance to data controllers in terms of the initial phase of the design of security measures, and abstain from imposing technology specific security requirements in order to grant flexibility to controllers to adapt their security mechanisms to their particular business model and the given data environment.

CHAPTER 9: Conclusion

In response to the rapidly changing landscape of emerging technologies that entail an increased level of re-identification risk, a shift in the data security regulation paradigm is required to overcome the shortcomings of conventional law-making and develop laws more easily adaptable to technological changes. This argument is summarised in the following words of the European Data Protection Supervisor (EDPS, 2015): *"Judging by the longevity of Directive 95/46/EC, it is reasonable to expect a similar timeframe before the next major revision of data protection rules, perhaps not until the late 2030s. Long before this time, data-driven technologies can be expected to have converged with artificial intelligence, natural language processing and biometric systems, empowering applications with machine-learning ability for advanced intelligence".* The key challenge is not only to set out the key objectives of a regulation, which is undoubtedly an important part of the entire law-making process, but mainly to define and operate regulation able to realise the governmental goals set in the most effective and efficient way possible. In this regard, the existing regulatory approach to data security adopted by the EU legislator is falling short of achieving the goals initially defined as it fails to effectively address the problem of cyber insecurity. In order for regulation to innovate and transform in a manner akin to the way industry evolves, and be agile in terms of both the processes and means of achieving regulatory objectives, regulators can leverage decision-making industry practices. It should be recognised that the regulatory environment differs from that of a firm in many aspects and therefore it would be politically and socially unacceptable to experiment on a market-wide scale, where a mistake is magnified many times over. Yet, this does not mean that regulators should abstain from adopting new approaches to regulation but, instead, they should find a way of doing so that increases market confidence and the quality of intervention by adopting an incentive-based approach to compliance able to meet the needs of the current complex and fast-paced business environments. To this end, the inclusion of intervention agents and, especially, the collaboration of technical and policy experts from government, industry, academia, non-governmental and consumer groups in the early stages of the law-making process, and throughout the process, can ensure that regulators are not left behind from technical developments in industry since it helps them overcome the problems stemming from the lack of technical knowledge. In this context, the role of government must principally be that of a catalyst or a facilitator which enables a coordinated collaboration between surrogate regulators by filling any gaps that may exist and facilitating links between the different layers

of the regulatory process by creating the necessary preconditions for stakeholders to assume a greater share of the regulatory burden.

The core of the protection-utility trade-off lies in finding ways to structure data so that meaningful information can be derived from it. In this respect, a regulatory framework needs to be built for balancing the potential benefits of data processing operations against the potential harms of particular practices. Such a framework may include a set of best practices for data controllers articulating in non-technical terms the guarantees provided by the various data protection mechanisms and determining the parameters to be considered in evaluating whether certain processing activities can provide sufficient privacy without considerably affecting the data structure. In the post-anonymisation era, the risk of re-identification can serve as a point of reference for regulators to set a threshold, below which personal data can be considered protected under given circumstances, and thus treated accordingly. Explicitly recognising that de-identification involves a broad spectrum of practices, each of which has different levels of strength, and identifying certain key points along that spectrum, which would establish a threshold, may have significant regulatory and policy implications as it enables the development of regulatory guidance that brings not only regulatory relief but also encourages the maximum use of de-identification compatible with the purposes of the data processing. Even though the revelation of the breach of purportedly anonymised datasets has resulted in both regulators and private actors reconsidering the effectiveness of the anonymisation technique, policy has not yet fully responded to the problem as the current EU approach still relies on the binary distinction between personal and anonymous data. The only way to reframe the debate over the feasibility of perfect anonymity and thus bridge the concerns of formalists and pragmatists lies in changing the focus of data protection law. In order to build a clear and workable legal framework, focus needs to be placed on the processes that decrease the likelihood of harm in the first place. Given that perfect data protection is not a realistic scenario, which means that there will always remain the risk of re-identification even in the cases when the risk is remote, efforts to prevent the harm have proved to be fruitless. It is high time data protection law changed its focus toward the process of risk management, that is, the procedures aiming to minimise the risk, which should be implemented ex ante. Approaches to data security that revolve around specific data subjects and the nature of information collected, used or disclosed, instead of taking into consideration datasets as a whole – the data environment - seem inadequate, at least by themselves, to respond to the failure of anonymisation. The same holds true with ex post, individualised

remedies based on the degree of harm caused; specific harms can be hard to articulate or even locate.

Based on the definition given by Bambauer (2012), according to which security comprises "the set of technological mechanisms (including, at times, physical ones) that mediates requests for access or control", the role of data security regulation is to address the selection and implementation of those mechanisms by determining who is able to access, use and alter data. In this regard, data security regulation is conceived as a general process which encompasses the process of constantly identifying risk, the process of developing and implementing administrative, legal, technical and physical safeguards in order to protect against data breaches, as well as the process of developing a response plan in the event of a data breach. In other words, the main precondition for a company to be held liable should not be associated with the causation of harm but rather with the failure to take the necessary steps to sufficiently reduce the risk; that is why, a company should be held liable even in the absence of a data breach (Hartzog & Solove, 2015). The actual harm caused is a relevant factor only insofar as it helps identify the procedures that might not have been properly implemented. Data security policy should be based on the idea that parties having custodian-like responsibilities ought to take steps to protect those who have entrusted them with their data (Rubinstein & Hartzog, 2015). Consequently, process failures or infringement of reasonableness security standards should be treated as the reasons for holding companies liable irrespective of the action of others or the causation of harm. The fact that data security regulation should be "agnostic about ex post harm in favour of ex ante controls" implicitly indicates that a certain degree of harm is acceptable; the processes implemented can only guarantee an acceptable level of data protection and hence data security policy can only be conceived of as risk tolerant (Ohm, 2010; Yakowitz, 2011; Wu, 2013; Rubinstein & Hartzog, 2015). The locus of data security policy should be the process of mitigating risks, a process driven by policies balancing data protection with data utility. Neither policymakers nor technologists alone can guarantee the protection of personal data. Data security policy requires a careful equilibrium on multiple fonts: law and technology, privacy and utility, data treatment and controls. The advances in the field of technology which allow for new uses of data have the potential to bring a wide range of benefits, as already witnessed in the case of Big data, but at the same time new types of harm can be generated by the various means of processing the data. Recognising the need to strike an appropriate balance between new

opportunities and individual values means that the appropriate security measures must be placed on data depending on the potential benefits and the created risks.

The proposed regulatory model aims at providing a clear framework to data controllers with regard to the assets that should be protected and the level of security required covering all the aspects of the security of personal data processing. The model adopts a hybrid approach to the regulation of data controllers' security obligations, which combines high-level security standards, or else the outcome to be achieved, with prescriptive guidance – examples of security measures that should be implemented in order to achieve the required outcome. By laying down high-level security objectives, which could reasonably be applied to a wide number of organisations handling personal data, and providing a baseline, that is, a minimum set of acceptable compliance strategies that must be in place, the regulatory model aspires to provide legal certainty without hampering innovation since no specific technological designs are mandated. The model establishes a layered approach, according to which data controllers should apply different levels of security – basic, medium, high - depending on certain risk factors. A layered approach to data security suggests that regulators establish a sufficiently comprehensive model appropriate for different contexts that would be regularly updated with the input of experts and stakeholders. The aim of the model is its components to specify a set of contexts in which certain security requirements would apply and case-by-case review would be avoided in the initial phase of the design of security measures. In addition, the regulatory model proposed introduces a dynamic approach to the required security level, which is contingent upon the data environment of each processing activity in the sense that there is no fixed level of the re-identification risk involved in each processing but the risk is rather defined by the interplay between the risk factors. The objective of the proposed model is to serve as a step forward to bridging the gap between the divergences in legal and technical perspectives as it requires lawmakers to be equipped with technical knowledge. In addition, the model can serve as an example of the way laws regulating the security obligations of data controllers should be drafted as it aspires to accomplish two goals – address the issues arising from the technology neutral approach to security obligations by providing clear guidance to data controllers in terms of the initial phase of the design of security measures, and abstain from imposing technology specific security requirements in order to grant flexibility to controllers to adapt their security mechanisms to their particular business model and the given data environment. Two things should be borne in mind with respect to this model. First, this model solely serves as an example of how data security

regulation should be formed. Second, it is essential that the model is regularly adapted to the technological changes not only in terms of the security objectives suggested but also with regard to the categories of data that should require special treatment. The same applies to the level of security required since what is currently considered to be a medium-level security measure might be a basic security measure in ten years' time.

The contribution the work presented in this thesis has made to the body of knowledge lies in the fact that it suggests a new approach to the data security regulation, which not only places the focal point to the ex-ante mitigation of the re-identification risk, as opposed to the current approach that imposes ex-post sanctions, but also introduces the industry's practices in the law-making process. In contrast to the previous work in this field, which either addresses a few of the aspects of the problem (e.g. only legal or technical) or makes recommendations that take into consideration only some of the factors affecting the re-identification risk level, this thesis provides a systematic assessment of the steps that should be taken by the private sector to determine the security level required for each data processing operation. In particular, the proposed model aims at providing clear guidance to both organisations and regulators as to the types of security measures that should be considered appropriate in each case by tailoring the security obligations of data controllers to the level of re-identification risk for different states of data, based on the interplay between the risk factors, and suggesting security measures proportionate to the re-identification risk level. In this respect, the contribution of this work also lies in bridging the gap between lawmakers, organisations and technologists as it not only examines the matter at stake from all three perspectives but it also takes a step further by integrating these perspectives and producing an innovative regulatory model. However, the scope of this thesis is limited to setting a theoretical ground as there is no empirical content included. Hence, the possible future directions of research coming out of this work would involve identifying the weaknesses of this model and filling the gaps by validating its viability on practical grounds. In particular, qualitative and quantitative methods can be used to estimate whether the private sector would find useful the guidance provided in determining the security measures that would best suit each processing activity. In addition, in order for this model to work in practice, research is required to explore the means of applying this theoretical framework to the law-making process and, more specifically, the changes that need to be made in the current legal construction to be able to embrace the innovative recommendations made in this thesis.

Bibliography

ABC4Trust. 2013. Position on the eIDAS Regulation. Seventh Framework Programme. Available at: https://abc4trust.eu/download/flyer/eIDAS_Final.pdf.

Abuhmed, T., Mohaisen, A. and Nyang, D. 2008. A Survey on Deep Packet Inspection for Intrusion Detection Systems. March. Available at: http://arxiv.org/pdf/0803.0037.pdf.

Acquisti, A. and Gross, R. 2009. Predicting Social Security Numbers From Public Data. Carnegie Mellon University, Pittsburgh. Available at: http://www.pnas.org/content/106/27/10975.full.pdf.

Agencia Española de Protección de Datos. 2013. Guía para clientes que contraten servicios de cloud computing. Available at : https://www.agpd.es/portalwebAGPD/canaldocumentacion/publicaciones/common/Guias/GUIA_Cloud.pdf.

Allan, C. 2007. Adaptive management of natural resources. Proceedings of the 5th Australian Stream Management Conference. Australian rivers: making a difference. Charles Sturt University. Thurgoona, South Wales. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.705.5144&rep=rep1&type=pdf.

Allot Communications Ltd. 2007. White Paper - Digging Deeper into Deep Packet Inspection. Available at: http://www.datacom.cz/files_datacom/dpi_white_paper.pdf.

Alpcan, T. and Basar, T. 2011. *Network Security: A Decision and Game-Theoretic Approach.* Cambridge: Cambridge University Press.

Anderson, D. 2015. A Question of Trust: Report of the Investigatory Powers Review. Presented to the Prime Minister pursuant to section 7 of the Data Retention and Investigatory Powers Act 2014. June 2015. Available at: https://terrorismlegislationreviewer.independent.gov.uk/wp-content/uploads/2015/06/IPRReport-Print-Version.pdf.

Anderson, N. 2007. Deep packet inspection meets 'Net Neutrality, CALEA. *ArsTechnica*. July 24. Available at: http://arstechnica.com/gadgets/2007/07/deep-packet- inspectionmeets-net-neutrality/.

Andronikou, V., Yannopoulos, A. and Varvarigou, T. 2008. Biometric Profiling: Opportunities and Risks. In Hildebrandt, M. and Gutwirth, S. (eds.). *Profiling the EuropeanCitizen: Cross-disciplinary Perspectives*. Springer Science and Business Media B.V. pp. 131-146.

Antonello, R., Fernandesa, S., Kamienskib, C., Sadoka, D., Kelnera, J., Gódorc, I., Szabóc,G. and Westholmd, T. 2012. Deep packet inspection tools and techniques in commodityplatforms: Challenges and trends. *Journal of Network and Computer Applications* 35(6):1863-1878.

Arrington, M. 2006. AOL proudly releases massive amounts of user search data. *TechCrunch.* Available at: http://tinyurl.com/AOL-SEARCH-BREACH.

Artan, N.S. and Chao, H.J. 2007. Design and analysis of a multipacket signature detection system. *International Journal of Security and Networks* 2(1–2): 122-136.

Article 29 Data Protection Working Party. 2005. Working Document on Data Protection Issues Related to RFID Technology. WP 105. January 19. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2005/wp105_en.pdf.

Article 29 Data Protection Working Party. 2007. Opinion 4/2007 on the Concept of Personal Data. 01248/07/EN WP 136. June 20. Available at: http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2007/wp136_en.pdf.

Article 29 Data Protection Working Party. 2008. Opinion 1/2008 on Data Protection Issues Relating to Search Engines. 00737/EN WP 148. April 4. Available at: http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2007/wp136_en.pdf.

Article 29 Data Protection Working Party. 2009. Opinion 5/2009 on Online Social Neworking. WP 163. June 12. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2009/wp163_en.pdf.

Article 29 Data Protection Working Party. 2010. Opinion 1/2010 on the concept of ‘controller’ and ‘processor’. 00264/10/EN WP169. Adopted on 16 February. Available at: http://tinyurl.com/WP29-CONT-PROC.

Article 29 Data Protection Working Party. 2011. Opinion 13/2011 on Geolocation Services on Smart Mobile Devices. WP 185. May 16. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinionrecommendation/files/2011/wp185_en.pdf.

Article 29 Data Protection Working Party. 2012. Opinion 06/2012 on the draft CommissionDecision on the measures applicable to the notification of personal data breachesunder
Directive 2002/58/EC on privacy and electronic communications. 01119/13/EN WP197. Adopted on 12 July. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp197_en.pdf.

Article 29 Data Protection Working Party. 2014a. Opinion 05/2014 on Anonymisation Techniques. Adopted on 10 April. Available at: http://ec.europa.eu/justice/dataprotection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

Article 29 Data Protection Working Party. 2014b. Statement on the Role of a Risk-Based Approach in Data Protection Legal Frameworks. WP 218. Adopted on 30 May 2014. Available at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp218_en.pdf.

Article 29 Data Protection Working Party & Working Party on Police and Justice. 2009. TheFuture of Privacy. Joint Contribution to the Consultation of the European Commission onthe Legal Framework for the Fundamental Right to Protection of Personal Data. WP 168. Available at: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2009/wp168_en.pdf.

Ashcroft, J. 2001. Testimony of the Attorney General to the Senate Committee on the Judiciary. Washington D.C. September 25. Available at: https://www.justice.gov/ag/9-25-01-attorney-general-john-ashcroft-testimony-senate-committee-judiciary.

AustralianGovernment.2009.Cyber Security Strategy. Available at: https://www.ag.gov.au/RightsAndProtections/CyberSecurity/Documents/AG%20Cyber%20Security%20Strategy%20-%20for%20website.pdf.

Australian Government. 2011. Connecting with Confidence: Optimising Australia's Digital Future.Available at: http://cyberwhitepaper.dpmc.gov.au/sites/default/files/documents/connecting_with_confidence_public_discussion_paper.pdf.

Australian Government. 2013. Big Data Strategy - Issues Paper. Department of Finance and Deregulation. March 2013. Available at: https://www.finance.gov.au/files/2013/03/Big-Data-Strategy-Issues-Paper1.pdf.

Ayres, I. and Braithwaite, J. 1992. *Responsive Regulation: Transcending the Deregulation Debate*. Oxford Socio-Legal Studies. Oxford: Oxford University Press.

Bahadur, G., Chan, W. and Weber, C.H. 2002. *Privacy Defended: Protecting Yourself Online*. Indianapolis: Que Corp.

Baker, F. and Savola, P. 2004. Ingress Filtering for Multihomed Networks. Network WorkingGroup. Available at: https://tools.ietf.org/html/bcp84#page-4.

Baker, S. and Yeo, M. 1999. Survey of Electronic and Digital Signature Initiatives. *InternetLaw & Technology Forum*. Available at: http://www.ilpf.org/groups/survey.htmI.

Balkin, J.M. and Levinson, S. 2006. The Processes of Constitutional Change: From PartisanEntrenchment to the National Surveillance State. *FORDHAM L. REV*. 75:489.

Bambauer, D.E. 2012. The Myth of Perfection, 2 *WAKE FOREST L. REV. ONLINE* 22. Available at: http://wakeforestlawreview.com/2012/04/the-myth-of-perfection/.

Barbaro, M. and Zeller, T. 2009. A Face Is Exposed for AOL Searcher No. 4417749. *N.Y. TIMES*.August 9.Available at: http://www.nytimes.com/2006/08/09/technology/09aol.html?_r=0.

Barofsky, A. 2000. The European Commission's Directive on Electronic Signatures: Technological "Favoritism" Towards Digital Signatures. *B.C.Int'l & Comp. L. Rev.* 24 (1):145-159.

Basalisco, B., Reid, A. and Richards, P. 2010. Interdependent Innovation in Telecommunications: Risk, Standardization and Regulation. In A. Gentzoglanis & A.

Henten (eds.). *Regulation and the Evolution of the Global Telecommunications Industry*. UK: Edward Elgar Publishing Inc. pp. 275-300.

Bateson, N. 1984. *Data Construction in Social Surveys*. London: George Allen and Unwin.

Battelle, J. 2006. *The Search: How Google and its rivals rewrote the rules of business and transformed our culture.* New York: Penguin Books Ltd.

Boyd, D. and Crawford, K. 2012. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society* 15(5): 662-679. Available at: http://dx.doi.org/10.1080/1369118X.2012.678878.

Bendrath, D. 2009. Global technology trends and national regulation: Explaining Variation inthe Governance of Deep Packet Inspection. Paper prepared for the International Studies Annual Convention New York City, 15-18 February 2009. Available at: http://userpage.fuberlin.de/~bendrath/Paper_Ralf-Bendrath_DPI_v1 - 5.pdf.

Bendrath, R. & Mueller, M. 2011. The end of the net as we know it? Deep packet inspection and Internet governance. *New Media & Society* 13(7): 1147.

Besanko, D. 1987. Performance versus Design Standards in the Regulation of Pollution. *Journal of Public Economics* 34(1): 19-44.

Bianchi, C. 2016. *Dynamic Performance Management*. Switzerland: Springer International Publishing.

Biddle, C.B. 1997. Legislating Market Winners: Digital Signature Laws and the Electronic Commerce Marketplace. *SAN DIEGO L. REV.* 34:1225.

Birnhack, M. 2013. Reverse Engineering Informational Privacy Law. *Yale Journal of Law and Technology* 15(1): 23-91.

Blythe, S.E. 2005. Digital Signature Law of the United Nations, European Union, United Kingdom and United States: Promotion of Growth in E-Commerce with Enhanced Security. *RICH. J.L. & TECH*. 11(2).

BOND. 2016. Adaptive Management: What it means for CSOs. Available at: https://www.bond.org.uk/sites/default/files/resource-documents/adaptive_management_-_what_it_means_for_csos_0.pdfBoss, A.H. 1999. The Internet and the Law: Searching for Security in the Law of Electronic Commerce. *NOVA L. REV*. 23:585.

Boschi, E. 2010. Privacy, Data Protection Law and Flow Data: Requirements, Issues, and Challenges. Available at: http://www.cert.org/flocon/2008/presentations/Boschi_flocon08.pdf.

Bowman, D. 2009. Sandvine presentation to the Canadian Radio-television and Telecommunications Commission. CRTC Public Notice 2008-19, July 6. Available at: http://www.crtc.gc.ca/public/partvii/2008/8646/c12_200815400/1241688.DOC.

Boyd, J. 1987. Organic Design for Command and Control. Available at: https://danford.net/boyd/organic.pdf.

Brazell, L. 2016. Trust in electronic transactions and e-signatures: the new EU "eIDAS" regime. January 15. Available at: http://www.osborneclarke.com/connectedinsights/publications/trust-electronic-transactions-and-e-signatures-new-eu-eidas-regime/.

Bremler-Barr, A., Harchol, Y. and Hay, D. 2011. Space-time trade-offs in Software-based Deep Packet Inspection. IEEE International Conference on High Performance Switching and Routing (IEEE HPSR), Belgrade, Serbia. Available at: http://www.cs.huji.ac.il/~dhay/publications/BHH11.pdf.

Brenton, C. & Hunt, C. 2003. *Mastering Network Security* (2nd ed.). Alameda: SYBEX Inc.

Breyer, S. 1982. *Regulation and its Reform*. USA: Harvard University Press.

Bright, M. 2004. BT puts block on child porn sites. *The Guardian*. June 6. Available at:http://www.theguardian.com/technology/2004/jun/06/childrensservices.childprotection.

Brown, N. 2010. Legal Aspects of Information Security: Data Protection. Available at: https://neilzone.co.uk/masters/lais_theme_2_report.pdf.

Brownsword, R. 2008. *Rights, Regulation, and the Technological Revolution*. Oxford: Oxford University Press.

Brownsword, R. and Yeung, K. 2008. Regulating Technologies: Tools, Targets and Thematics. In R. Brownsword & K. Yeung (eds.). *Regulating Technologies: Legal Futures, RegulatoryFrames and Technological Fixes.* Oxford: Hart. pp. 3-22.

Brunton, F, and Nissenbaum, H. 2015. *Obfuscation: A User's Guide for Privacy and Protest*. Cambridge: MIT Press.

Burton, C. and Hoffman, S. 2015. Personal Data, Anonymization, and Pseudonymization in the EU.*The WSGR Data Advisor*. September 15. Available at: http://www.wsgrdataadvisor.com/2015/09/personal-data-anonymization-and-pseudonymization-in-the-eu/Bygrave, L. 2002. *Data Protection Law: Approaching Its Rationale, Logic and Limit*. The Hague: Kluwer Law International.

Bygrave, L.A. 2002. Privacy-enhancing technologies: Caught between a rock and the hard place. *Privacy Law and Policy Reporter* 9: 135–137.

Cavoukian, A. 2006. Creation of a Global Privacy Standard. Information and Privacy Commissioner of Ontario, Canada. November 2006. Available at: https://www.ipc.on.ca/images/resources/gps.pdf.

Cavoukian, A. 2008. Privacy and radical pragmatism: Change the paradigm. White Paper. Information and Privacy Commissioner of Ontario, Canada. August 8. Available at: https://www.ipc.on.ca/images/Resources/radicalpragmatism_238816250000.pdf.

Cavoukian, A. 2010. Privacy by Design: Achieving the Gold Standard in Data Protection forthe Smart Grid. Information and Privacy Commissioner of Ontario, Canada. June 2010. Available at: https://www.ipc.on.ca/images/resources/achieve-goldstnd.pdf.

Cavoukian, A. 2011a. Privacy by Design: The 7 Foundational Principles. Information and Privacy Commissioner of Ontario, Canada. August 2009, revised January 2011. Information and Privacy Commissioner of Ontario. Available at: http://www.ipc.on.ca/images/Resources/7 foundationalprinciples.pdf.

Cavoukian, A. 2011b. Privacy by Design in Law, Policy and Practice: A White Paper for Regulators, Decision-makers and Policy-makers. Information and Privacy Commissioner of Ontario, Canada. August 2011. Available at: https://privacybydesign.ca/content/uploads/2011/08/pbd-law-policy.pdf.

Cavoukian, A, and Castro, D. 2014. Big Data and Innovation: Setting the Record Straight: De-identification Does Work. Available at: http://www2.itif.org/2014-big-data-deidentification.pdf.

Cavoukian, A. and El Emam, K. 2011. Dispelling the Myths Surrounding De-identification: Anonymisation Remains a Strong Tool for Protecting Privacy. Available at: https://fpf.org/wp-content/uploads/2011/07/Dispelling%20the%20Myths%20Surrounding%20De-identification%20Anonymization%20Remains%20a%20Strong%20Tool%20for%20Protecting%20Privacy.pdf.

CERT. 2013a. DNS amplification attacks and open DNS resolvers. *CERT.be*, April 9. Available at: https://www.cert.be/pro/docs/dns-amplification-attacks-and-open-dnsresolvers.

CERT. 2013b. Lessons from the Stophaus/CloudFlare/SpamhausDDoS for ISPs. ComputerEmergency Response Team Australia. Available at: http://www.cert.at/services/blog/20130328190708-815_en.html.

Chen, N. Chen, X., Xiong, B., Lu, H. 2010. Method for tracing back abnormal packets in network transport layer. *Journal of Huazhong University of Science and Technology* (Nature Science Edition) 38(1): 9-13.

Cheng, T.H., Lin, Y.D., Lai, Y.C. and Lin, P.C. 2011. Evasion Techniques: Sneaking throughYour Intrusion Detection/Prevention Systems. IEEE. Available at:

http://speed.cs.nctu.edu.tw/~ydlin/pdf/Evasion_Techniques_Sneaking_through_Your_Intrusion_Detection_Prevention_Systems.pdf.

Chirumalla, K. 2017. Clarifying the feedback loop concept for innovation capability: A literature review. Presented at The XXVIII ISPIM Innovation Conference – Composing the Innovation Symphony, Austria, Vienna on 18-21 June 2017. https://www.researchgate.net/publication/316857374_Clarifying_the_feedback_loop_concept_for_innovation_capability_A_literature_review.

CISCO. 2006. Working with IP Addresses. *The Internet Protocol Journal* 9(1). Available at:http://www.cisco.com/c/en/us/about/press/internet-protocol-journal/back-issues/tablecontents-12/ip-addresses.html.

CISCO Support Community.2014. ISP Reporting Open DNS Resolvers. Available at: https://supportforums.cisco.com/thread/2260378.

Citron, D.K. 2007. Technological Due Process. *Washington University Law Review* 85: 1249-1313.

Clarke, R. and Knake, R. 2010. *Cyber War: The Next Threat to National Security and Whatto Do About It.* New York: HarperCollins.

Clayton, R. 2005. Failures in a Hybrid Content Blocking System. In G. Danezis& D. Martin(eds.). *Privacy Enhancing Technologies 5th International Workshop Revised Selected*
Papers. Cavtat. Croatia, May/June 2005. pp. 78-92.

Clayton, R. 2008. The Phorm "Webwise" system. May 18. http://www.cl.cam.ac.uk/~rnc1/080518-phorm.pdf.

Clayton, R. 2009. IWF, Wikipedia and the "Wayback Machine". UKNOF13, Sheffield, May28. Available at: http://www.uknof.org.uk/uknof13/Clayton-IWF.pdf.

CNIL. 2010. Security of Personal Data. Available at: http://www.cnil.fr/fileadmin/documents/en/Guide_Security_of_Personal_Data-2010.pdf.

CNIL. 2012a. Measures for the Privacy Risk Treatment. Available at: http://www.cnil.fr/fileadmin/documents/en/CNIL-ManagingPrivacyRisks-Measures.pdf.

CNIL. 2012b. Methodology for Privacy Risk Management: How to Implement the Data Protection Act. June 2012. Available at: https://www.cnil.fr/sites/default/files/typo/document/CNIL-ManagingPrivacyRisks-Methodology.pdf.

CNN Money. 2010. 5 Data Breaches: From Embarrassing To Deadly. Available at:http://tinyurl.com/CNN-BREACHES/.

Coglianese, C. and Lazer, D. 2003. Management-Based Regulation: Prescribing Private Management to Achieve Public Goals. *Law & Society Review* 37: 691.

Coglianese, C., Nash, J. and Olmstead, T. 2003. Performance-Based Regulation: Prospects and Limitations in Health, Safety, and Environmental Protection. *Administrative Law Review* 55(4): 705-729.

Cohen, T. 2016. 21st Century Regulation Putting Innovation at the heart of Payments Regulation. PayPal – ebay Inc. Available at: https://www.paypalobjects.com/webstatic/en_US/mktg/public-policy/PayPal-Payment-Regulations-Booklet-US.pdf.

Comer, D.E. 2006. *Interworking with TCP/IP: Principles, Protocols and Architectures* (5thed.). New Jersey: Prentice Hall.

Commission of the European Communities. 2003. First Report on the Implementation of the Data Protection Directive (95/46/EC). COM (2003) 265final. May 15. Available at: http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2003:0265:FIN:EN:PDF.

Committee on Automated Personal Data Systems. July 1973. Available at: https://epic.org/privacy/hew1973report/.

Cooper, A. 2011. Doing the DPI Dance: Assessing the Privacy Impact of Deep Packet Inspection. In W.Aspray& P. Dotty (eds.). *Privacy in America: Interdisciplinary Perspectives.* Maryland: Scarecrow Press. pp. 139-166.

Corwin, P.S. 1998. Digital Signatures and Signature Dynamics: Some Issues to Consider. *BANKING POL'y REP*. 17:9.

Costa, L. and Poullet, Y. 2012. Privacy and the regulation of 2012. *Computer Law and Security Review* 28(3): 254-262.

Council of Europe. 2001. Convention on Cybercrime Explanatory Report. Adopted on November 8. Available at: http://conventions.coe.int/.

Crawford, K. and Schultz, J. 2014. Big data and due process: Toward a framework to redress predictive privacy harms. *BCL Rev* 55(1): 93. Available at: http://tinyurl.com/BD-HARMS.

Dale, B.G., van der Wiele, T. and van Iwardeen, J. 2007. *Managing Quality* (5th ed.). Oxford: Blackwell Publishing.

Daly, A. 2010. The legality of deep packet inspection. *International Journal Commercial law& Policy*, 14(1).

Daly J. and Foushee, M. 1981. Performance Standards: A Practical Guide To The Use of Performance Standards As A Regulatory Alternative. Project on Alternative Regulatory Approaches. Available at: http://morganrobertson.com/Wetlandia%20Files/s1%201981%20PARA%20Perf%20Standards.pdf.

Damas, J.L. &Karrenberg, D. 2008. Network Hygiene Pays Off: The Business Case for IP Source Address Verification. Available at: ftp://ftp.ripe.net/ripe/docs/ripe-432.pdf.

Danagher, L. 2012. An Assessment of the Draft Data Protection Regulation: Does it Effectively Protect Data?. *European Journal of Law and Technology* 3(3). Available at: http://ejlt.org/article/view/171/260.

Davies, C.J. 2009. The hidden censors of the Internet. *Wired*, May 20. Available at: http://www.wired.co.uk/wired-magazine/archive/2009/05/features/the-hidden-censors-ofthe-internet.aspx?page=all.

De Montjoye, Y.A., Hidalgo, C.A., Verleysen, M. and Blondel, V.D. 2013. Unique in the Crowd: The Privacy Bounds of Human Mobility. *Scientific Reports*3: 1376. Available at: https://www.nature.com/articles/srep01376.

De Montjoye, Y.A., Radaelli, L. and Singh, V.K. 2015. Unique in the shopping mall: On the re-identifiability of credit card metadata. *Science* 347(6221): 536-539. Available at: http://tinyurl.com/UNIQ-CC.

Der Hessische Datenschutzbeauftragte. 2015. Key data protection points for the trilogue on the General Data Protection Regulation. Conference of the Data Protection Commissioners of the Federal Government and the Federal States. 14 August 2015.Availableat:https://www.democraticmedia.org/content/key-data-protection-points-trilogue-general-data-protection-regulation.

De Visser, M. 2009. *Network-Based Governance in EC Law: The Example of EC Competition and EC Communications Law*. Oxford: Hart Publishing.

Derner, J.D. and Augustine, D.J. 2016. Adaptive Management for Drought on Rangelands, *Rangelands* 38(4): 211-215.

Dilanian, K. 2012. U.S. Chamber of Commerce leads defeat of cyber-security bill. *Los Angeles Times*. August 3. Available at: http://articles.latimes.com/2012/aug/03/nation/la-nacyber-security-20120803.

Domingo-Ferrer, J. and Torra, V. 2008. A critique of k-anonymity and some of its enhancements. In *3rd Intl. Conference on Availability, Reliability and Security (ARES 2008)*, Los Alamitos CA: IEEE Computer Society, 2008: 990-993, DOI: 10.1109/ARES.2008.97.

Domingo-Ferrer, J., Snachez, D. and Soria-Comas, J. 2016. *Database Anonymization: Privacy Models, Data Utility, and Microaggregation-based Intermodel*

Connections; Synthesis Lectures on Information Security, Privacy, & Trust 15: Morgan & Claypool, DOI: 10.2200/S00690ED1V01Y201512SPT015.

Douligeris C. and Mitrokotsa A. 2004. DDoS attacks and defence mechanisms: classificationand state-of-the-art. *Computer Networks* (44), 643-666, Elsevier.

Douligeris C. and Serpanos D. 2006. Network Security: Current Status and Future Directions.Wiley – IEEE.

Drucker, P.F. 1967. Technological trends in the twentieth century. In M. Kranzberg & C.W.Pursell (eds.).Technology in Western Civilization. New York: Oxford University Press. pp.10-21.

Dupont, B. 2013. Cybersecurity Futures: How Can We Regulate Emergent Risks?. *Technology Innovation Management Review*. Available at: http://timreview.ca/sites/default/files/article_PDF/Dupont_TIMReview_July2013.pdf.

Dutch Ministry of Security and Justice. 2011. Strength through cooperation. The National Cyber Security Strategy (NCSS). Available at: http://english.nctb.nl/Images/cyber-securitystrategy-uk_tcm92- 379999.pdf.

Dwork, C. and Pottenger, R. 2013. Towards Practicing Privacy. 20 *J. AM. MED. INFORMATICS ASS'N* 102.

EDPS. 2015. Annex to Opinion 3/2015: Comparative table of GDPR texts with EDPS recommendations. Opinion 3/2015: Europe's big opportunity - EDPS recommendations on the EU's options for data protection reform. July 27. Available at: https://edps.europa.eu/sites/edp/files/publication/15-07-27_gdpr_recommendations_annex_en_1.pdf.

Edwards, L. and Brown, I. 2009. Data Control and Social Networks: Irreconcilable Ideas?. 2009. In Andrea M. Matwyshyn (ed.). 2009. Harboring Data: Information Security, Law And The Corporation. Stanford: Stanford University Press. pp. 202-207.

Electronic Commerce Expert Group. 1998. Electronic Commerce: Building the Legal Framework. Report to the Attorney General of Australia. March 31. Available at: http://law.gov.au/aghome/advisory/eceg/ecegreport.html.

El Emam, K. 2010. Risk-Based De-Identification of Health Data. IEEE Security and Privacy 8: 64-67. Available at: http://www.ehealthinformation.ca/wp-content/uploads/2014/08/2010-Risk-based-de-identification-of-health-data.pdf.

El Emam, K. 2013a. *Guide to the De-Identification of Personal Health Information*. New York: CRC Press.

El Emam, K. 2013b. *Risky Business: Sharing Health Data while Protecting Privacy*. Bloomington, Indiana: Trafford Publishing.

El Emam, K. and Arbuckle, L. 2014a. *Anonymizing Health Data (*2nd ed.). Sebastapol, California: O'Reilly media.

El Emam, K. and Arbuckle, L. 2014b. De-Identification: A Critical Debate. Available at: https://fpf.org/2014/07/24/de-identification-a-critical-debate/.

El Emam, K., Gratton, E., Polonetsky, J. and Arbuckle, L. 2016. The Seven States of Data: When is Pseudonymous Data Not Personal Information?. Brussels Privacy Symposium on Identifiability: Policy and Practical Solutions for Anonymisation and Pseudonymisation. November 8. Brussels, Belgium. Available at: https://fpf.org/wp-content/uploads/2016/11/El-Emam_States-of-Data-Main-Article-short-v6.pdf.

El Emam, K., Jonker, E., Arbuckle, L., and Malin, B. 2015. A Systematic Review of Re-Identification Attacks on Health Data. *PLoS ONE* 10(4): e0126772. Available at: https://doi.org/10.1371/journal.pone.0126772.

Elliot, M. J. 1996. Attacks on Confidentiality Using the Samples of Anonymised Records; In *Proceedings of the Third International Seminar on Statistical Confidentiality*. Bled, Slovenia, October 1996. Ljubljana: Statistics SloveniaEurostat.

Elliot, M. J., Dibben, C., Gowans, H., Mackey, E., Lightfoot, D., O'Hara, K. and Purdam, K. 2015. Functional Anonymisation: The crucial role of the data environment in determining

the classification of data as (non-) personal. CMIST work paper 2015-2. Available at: http://tinyurl.com/FUNCANON.

Elliot, M. J., Mackey, E., O'Hara, K. and Tudor, C. 2016. *The Anonymisation Decision-Making Framework.* University of Manchester: UKAN.

Emam, K.E., and Alvarez, C. 2014. A Critical Appraisal of the Article 29 Working Party Opinion 05/2014 on Data Anonymisation Techniques. Oxford Journals – International PrivacyLaw. Available at: http://idpl.oxfordjournals.org/content/early/2014/12/12/idpl.ipu033.full.pdf+html.

ENISA. 2011a. Cyber security: future challenges and opportunities. Available at: www.enisa.europa.eu/publications/position-papers/cyber-security-future-challenges-andopportunities.

ENISA. 2011b. Technical Guideline for Minimum Security Measures: Guidance on the security measures in Article 13a. Version 1.0, December 2011. Available at: https://resilience.enisa.europa.eu/article-13/guideline-for-minimum-securitymeasures/technical-guideline-for-minimum-security-measures-v1.0.

ENISA. 2012a. Cyber Incident Reporting in the EU: An overview of security articles in EUlegislation. August 2012. Available at: http://www.enisa.europa.eu/activities/Resilienceand-CIIP/Incidents-reporting/cyber-incident-reporting-in-the-eu.

ENISA. 2012b. National Cyber Security Strategies. Setting the course for national efforts tostrengthen security in cyberspace. Available at: www.enisa.europa.eu/activities/Resilienceand-CIIP/national-cyber- security-strategies-ncsss/cyber-security-strategies-paper.

ENISA. 2012c. Study on data collection and storage in the EU. February 23. Available at: https://www.enisa.europa.eu/publications/data-collection.

ENISA. 2013a. Internet Service Providers fail to apply filters against big cyber- attacks. European Network and Information Security Agency, April 12. Available at:

http://www.enisa.europa.eu/media/press-releases/eu-agency-enisa-internet-serviceproviders-fail-to-apply-filters-against-big-cyber-attacks.

ENISA. 2013b. Recommended cryptographic measures: Securing personal data. September20. Available at: https://www.enisa.europa.eu/activities/identity-andtrust/library/deliverables/recommended-cryptographic-measures-securing-personal-data.

ENISA. 2014a. Privacy and Data Protection by Design – from policy to engineering. December 2014. Available at: https://www.enisa.europa.eu/activities/identity-andtrust/library/deliverables/privacy-and-data-protection-by-design.

ENISA. 2014b. Technical Guideline on Security Measures: Technical guidance on the security measures in Article 13a. Version 2.0, October 2014. Available at: http://www.enisa.europa.eu/activities/Resilience-and-CIIP/Incidents-reporting/technicalguideline-on-minimum-security-measures.

ENISA. 2015. Privay by Design In Big data. December 17. Available at: https://www.enisa.europa.eu/publications/big-data-protection.

Esayas S. Y. 2015. The role of anonymisation and pseudonymisation under the EU data privacy rules: beyond the 'all or nothing' approach. *European Journal of Law and Technology* 6(2).

Escudero A. 2001. Location data is as sensitive as content data. Contribution to the EU Forum on Cybercrime in Brussels. November 27. Available at: cybercrime-forum.jrc.it.

Escudero-Pascual, A. and Hosein, I. 2002. Understanding Traffic Data and Deconstructing Technology-neutral Regulations. March 7. Available at: http://www.it46.se/docs/papers/unece-latest-escuderoa-hoseini.pdf.

Escudero-Pascual, A. and Hosein, I. 2004a. The hazards of technology-neutral policy: questioning lawful access to traffic data. *Communications of the Association for Computer Machinery (CACM) Journal* 47(3). Available at: http://www.it46.se/docs/papers/acm-1905_prepub.pdf.

Escudero-Pascual, A. and Hosein, I. 2004b. The Hazards of Technology-Neutral Policy: Questioning Lawful Access to Traffic Data. Communications of the ACM, 47(3), 77–84.

e-Soft (2013). White Paper - Modern Network Security: The Migration to Deep Packet Inspection. December 16. Available at: http://esoft.untangle.com/content/pdf/dpi-migrationwhitepaper.pdf.

Etzioni, A. 2011.Cybersecurity in the Private Sector. *Issues in Science and Technology* 28, no. 1 (Fall 2011): 58-62.

Etzioni, A. 2013. The Bankruptcy of Liberalism and Conservatism. *Political Science Quarterly* 128(1): 39-65.

Etzioni, A. 2014. The Private Sector: A Reluctant Partner in Cybersecurity. *Georgetown Journal of International Affairs, International Engagement on Cyber* IV (3): 69-73.

Etzioni, A. 2015. *Privacy in a Cyber Age: Policy and Practice*. New York: Palgrave Macmillan.

EURIM. 2010. Can Society Afford To Rely on Security by Afterthought Not Design?. Statusreport and recommendations of the Information Society Alliance (EURIM) Subgroup onSecurity by Design. October 2010. Available at: http://www.eurim.org.uk/activities/ig/1010-SbD_Full.pdf.

EURIM. 2013. Position Paper on The Proposed General Regulation on Data Protection. Digital Policy Alliance. March 2013. Available at: http://dpalliance.org.uk/wpcontent/uploads/2013/03/1303_Data-Protection-Position-Paper2.pdf.

European Commission. 1999. Towards a new framework for Electronic Communications Infrastructure and Associated Services. The 1999 Communications Review. Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions. COM (1999) 539 Final. 10 November 1999. Available at: http://europa.eu/legislation_summaries/internal_market/single_market_services/l24216_en.htm.

European Commission. 2009. Telecoms: Commission launches case against UK over privacyand personal data protection. EC IP/09/570, April 14. Available at: http://europa.eu/rapid/press-release_IP-09-570_en.htm.

European Commission. 2010a. A comprehensive approach on personal data protection in theEuropean Union. Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions. COM(2010) 609 final. November 4, 2010. Available at:http://ec.europa.eu/justice/news/consulting_public/0006/com_2010_609_en.pdf.

European Commission. 2010b. Communication From The Commission To The European Parliament, The Council, The European Economic And Social Committee And The Committee Of The Regions: Smart Regulation in the European Union. COM/2010/0543 final. Brussels, 8 October 2010. Available at: http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A52010DC0543.

European Commission. 2012. How Will the EU's Data Protection Reform Benefit EuropeanBusinesses?. Available at: http://ec.europa.eu/justice/dataprotection/document/review2012/factsheets/7_en.pdf.

European Commission. 2013. Cybersecurity Strategy of the European Union: An Open, Safeand Secure Cyberspace. February 7. Available at: http://ec.europa.eu/digitalagenda/en/news/eu-cybersecurity-plan-protect-open-internet-and-online-freedom-andopportunity-cyber-security.

European Commission. 2014. Q&A: Electronic Identification and Trust Services (eIDAS) Regulation. October 14. Available at: http://europa.eu/rapid/press-release_MEMO-14-586_en.htm.

European Commission. 2015. Communication From The Commission To The European Parliament, The Council, The European Economic And Social Committee And The Committee Of The Regions: Better regulation for better results - An EU agenda. COM(2015) 215 final. Strasbourg, 19 May 2015. Available at: http://eur-lex.europa.eu/legal-content/en/txt/?uri=celex%3a52015dc0215.

European Commission. 2017. Better Regulation Guidelines. Commission Staff Working Document SWD (2017) 350. Brussels, 7 July 2017. Available at: https://ec.europa.eu/info/sites/info/files/better-regulation-guidelines.pdf.

European Council. 2012. Evaluation of the Implementation of the Data Protection Directive – Annex 2. Available at: http://lobbyplag.eu/governments/assets/pdf_all/CD-all.pdf.

European Data Protection Supervisor. 2011. Opinion of the European Data Protection Supervisor on net neutrality, traffic management and the protection of privacy and personal data. October 7. Available at: http://berec.europa.eu/files/document_register/2012/8/BoR_PC04_%2811%29_20_edps.pdf

European Data Protection Supervisor. 2015. Opinion 3/2015 - EDPS recommendations on the EU's options for data protection reform. July 27. Available at: https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2015/15-07-27_GDPR_Recommendations_EN.pdf.

European Union Agency for Fundamental Rights. 2013. Handbook on European data protection law. Belgium: Council of Europe. Available at: http://fra.europa.eu/sites/default/files/fra-2014-handbook-data-protection-law-2nded_en.pdf.

Falzone, A. 2013.Regulation and Technology. *Harvard Journal of Law & Public Policy* 36(1):105-107.

Federal Ministry of the Interior. 2011. Cyber Security Strategy for Germany. Available at: www.cio.bund.de/SharedDocs/Publikationen/DE/StrategischeThemen/css_engl_download.pdf?__blob=publicationFile.

Feiler, L. 2010. New Approaches to Network and Information Security Regulation: The EUTelecoms Package. *Computer Law Review International* 11(2): 43-49.

Feiler, L. 2011. Information Security Law in the EU and the U.S.: A Risk-Based Assessmentof Regulatory Policies. A joint initiative of Stanford Law School and the Universityof
Vienna School of Law TTLF Working Papers No. 9. Available at:

https://www.law.stanford.edu/publications/information-security-law-in-the-eu-and-the-us-%E2%80%94-a-risk-based-assessment-of-regulatory-policies.

Fell, M. 2013. Manifesto for Smarter Intervention in Complex Systems. Available at: https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/discussions/Smarter_Intervention_In_Complex_Systems%20%282013%29.pdf

Ferguson, P. 2013. Open DNS Recursive Resolvers, DNS Amplification Attacks, and BCP38: What Are They, and Why Should You Care? Available at: http://www.maawg.org/system/files/Fergie_DNS_Open_Resolver_MAAWG_India_SANOG.pdf.

Ferguson, P. and Senie, D. 2000. Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing. Network Working Group. Available at: http://tools.ietf.org/html/bcp38.

Financial System Inquiry. 2014. Regulation in a digital environment. Available at: http://fsi.gov.au/publications/interim-report/09-technology/regulation-digital-environment/.

Finnie, G. 2009. ISP traffic management technologies: The state ofthe art. On behalf of theCanadian Radio Television and Telecommunications Commission (CRTC). Heavy Reading,January2009. Available at: http://www.crtc.gc.ca/PartVII/eng/2008/8646/ispfsi.htm#_toc219621630.

Fischer, S.F. 2001. Saving Rosencrantz and Guildenstern in a Virtual World? A ComparativeLook at Recent Global Electronic Signature Legislation. *B.U. J. SCI. & TECH. L.*7:229-242.

Fletcher, N. 2007. Challenges for regulating financial fraud in cyberspace. *Journal of Financial Crime* 14(2), 190-207.

Frieden, R. 2008. Internet Packet Sniffing and Its Impact on the Network Neutrality Debateand the Balance of Power between Intellectual Property Creators and Consumers. *FordhamIntellectual Property, Media and Entertainment Law Journal* 18(3): 634-661.

FTC. 2010. Protecting Consumer Privacy In An Era of Rapid Change: A Proposed Framework For Businesses And Policymakers. Preliminary FTC Staff Report. December 2010. Available at: https://www.ftc.gov/sites/default/files/documents/reports/federal-tradecommission-bureau-consumer-protection-preliminary-ftc-staff-report-protectingconsumer/101201privacyreport.pdf.

FTC. 2012. Protecting Consumer Privacy in an Era of Rapid Change. Available at: https://www.ftc.gov/sites/de-fault/files/documents/reports/federal-trade-commission-report-protecting-con-sumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.

Fuchs, C. 2013a. Implications of DPI Internet Surveillance for Society. The Privacy & Security Research Paper Series, PACT, Issue 1. Available at: http://www.projectpact.eu/privacy-security-research-paperseries/%231_Privacy_and_Security_Research_Paper_Series.pdf.

Fuchs, C. 2013b. Societal and Ideological Impacts of Deep Packet Inspection Internet Surveillance. *Information, Communication & Society* 16:8, 1328-1359.

Gallagher, S. 2012. Big brother on a budget: How internet surveillance became so cheap. *Wired*, August 29. Available at: http://www.wired.co.uk/news/archive/2012-08/29/dpiinternet-surveillance.

Gandomi, A. and Haider, M. 2015. Beyond the Hype: Big Data Concepts, Methods, and Analytics. *International Journal of Information Management* 35(2): 137–144.

Garfinkel, S.L. 2015. De-identification of Personal Information. National Institute of Standards & Technology. NISTIR 805. October 2015. Available at: http://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf.

Gattuso, J.L. 2012. Ensuring Cybersecurity: More Red Tape Is Not the Answer. June 5. Available at:http://www.heritage.org/research/reports/2012/06/cybersecurity-and-red-tapemore-regulations-not-the-answer.

Geere, D. 201 2. How deep packet inspection works. *WIRED*, April 27. Available at: http://www.wired.co.uk/news/archive/2012-04/27/how-deep-packet-inspection-works.

Gellman, R. 2009. Privacy in the Clouds: Risks to Privacy and Confidentiality from Cloud Computing. World Privacy Forum. February 23. Available at: http://www.worldprivacyforum.org/pdf/WPF_Cloud_Privacy_Report.pdf.

Gersen, J.E. 2007. Temporary Legislation. *U. CHI. L. REV*. 74:247.

Gervais, D.J. 2005. Towards a New Core International Copyright Norm: the Reverse ThreeStepTest. *Marquette Intellectual Property Law Review* 9(1). Spring 2005. Available at:http://ssrn.com/abstract=499924.

Ginsberg, J., Mohebbi, M.H., Patel, R.S., Brammer, L., Smolinski, M.S. and Brilliant, L. 2009. Detecting Influenza Epidemics Using Search Engine Query Data. 457 *Nature* 1012.

Government of Canada. 2002. Lawful Access - Consultation Document. Department of Justice, Industry Canada, Solicitor General Canada. August 25. Available at: http://www.justice.gc.ca/eng/cons/la-al/la-al.pdf.

Government of Canada. 2010. Canada's Cyber Security Strategy: For a Stronger and More Prosperous Canada. Available at: www.publicsafety.gc.ca/prg/ns/cbr/ccss-scc-eng.aspx.

Gratton, E.2013. *Understanding Personal Information: Managing Privacy Risks*. Canada: LexisNexis.

Gratton, E. 2014. If Personal Information is Privacy's Gatekeeper, Then Risk of Harm is the Key: A Proposed Method for Determining What Counts as Personal Information. *ALB. L.J. SCI. & TECH*. 24: 105.

Greene, B.R & Smith, P. 2002. Cisco ISP Essentials. USA: Cisco Press.

Grimmelmann, J. 2009. Saving Facebook. *Iowa Law Review* 94:1137.

Gross, G. 2012. ISPs: No New Cybersecurity Regulations Needed. *PCWorld*. Available at: http://www.pcworld.com/article/251444/isps_no_new_cybersecurity_regulations_needed.html.

Gunningham, N. and Sinclair, D. 1998. Designing Smart Regulation, Available at: https://www.oecd.org/env/outreach/33947759.pdf.

Gürses, S. 2014. Can you engineer privacy? *Communications of the ACM* 57(8):20–23.

Gürses, S., Troncoso, C. and Díaz, C. 2011. Engineering Privacy by Design. Available at: http://boemund.dagstuhl.de/mat//Files/11/11061/11061.DiazClaudia.Paper.pdf.

Gutwirth, S., De Hert, P. and De Sutter, L. 2008. The trouble with technology regulation from a legal perspective: Why Lessig's 'optimal mix' will not work. In R. Brownsword & K. Yeung (eds.). *Regulating Technologies: Legal Futures, Regulatory Frames and Technological Fixes.* Oxford: Hart. pp. 193-218.

Gutwirth, S., Leenes, R. and De Hert, P. 2015. *Reforming European Data Protection Law*. Law, Governance and Technology Series, Issues in Privacy and Data Protection. Published by Springer. Available at: http://www.springer.com/gb/book/9789401793841.

Gutwirth, S., Leenes, R. and De Hert, P. 2016. *Data Protection on the Move: Current Developments in ICT and Privacy/Data Protection* (Vol. 24). Heidelberg: Springer.

Halftech, G. 2008. Legislative Threats. *Stanford Law Review* 61(3): 629-710.

Hammersely, B. 2012. *64 Things you Need to Know Now for Then: How to Face the Digital FutureWithout Fear.* London: Hodder and Stoughton.

Hansen, M. 2013. Data protection by default in identity-related applications. IDMAN 2013,volume 396 of IFIP AICT, pp. 4–17. IFIP International Federation for Information Processing, Springer.

Harrington, J. 2005. *Network Security: A Practical Approach*. San Fransisco: Elsevier Inc.

Hartzog, W. and Solove, D.J. The Scope and Potential of FTC Data Protection. 83 *GEO. WASH. L. REV.* 2230.

Hartzog, W. and Stutzman, F. 2013. The Case for Online Obscurity, 101 *CAL. L. REV.* 1.

Hemenway, D. 1980. Performance vs. Design Standards. Prepared for Office of Standards Information, Analysis Development, Office of Engineering Standards and National Bureau of Standards, U.S. Department of Commerce. Available at: http://gsi.nist.gov/global/docs/pubs/NISTGCR_80-287.pdf.

Heron, S. 2009. Online privacy and browser security. Network Security, Volume June (6): 4-7.

Hildebrandt, M. 2011. Legal Protection by Design: Objections and Refutations. Legisprudence5.2: 223-248. Available at: http://works.bepress.com/mireille_hildebrandt/43/.

Hildebrandt, M. 2013. Legal Protection by Design in the Smart Grid. Available at: http://works.bepress.com/mireille_hildebrandt/42.

Hildebrandt, M. and Tielemans, L. 2013. Data protection by design and technology neutral Law. *Computer Law & Security Review* 29: 509-521.

Hintze, M. Viewing the GDPR Through a De-Identification Lens: A Tool for Clarification and Compliance. Brussels Privacy Symposium on Identifiability: Policy and Practical Solutions for Anonymisation and Pseudonymisation. November 8. Brussels, Belgium. Available at: https://fpf.org/wp-content/uploads/2016/11/M-Hintze-GDPR-Through-the-De-Identification-Lens-31-Oct-2016-002.pdf.

Hon, W.K., Kosta, E., Millard, C., and Stefanatou, D. 2014. Cloud Accountability: The Likely Impact of the Proposed EU Data Protection Regulation. *Tilburg Law School Legal Studies Research Paper Series* No. 07/2014. Available at: http://ssrn.com/abstract=2405971.

House of Commons Culture, Media and Sport Committee. 2007-2008. Harmful content on the Internet and in video games. Tenth Report of Session 2007–08, Volume II. Available at: http://www.publications.parliament.uk/pa/cm200708/cmselect/cmcumeds/353/353ii.pdf.

House of Lords. 2014. Data Retention and Investigatory Powers Bill – Second Reading. July16. Available at http://www.theyworkforyou.com/lords/?id=2014-07-16a.599.10.

Hutty, M. 2004. Cleanfeed: the facts. *LINX Public Affairs*. Available at: https://publicaffairs.linx.net/news/?p=154.

House of Lords & House of Commons Joint Committee on Draft Communications Data Bill.2013. Draft Communications Data Bill. Written Evidence. Session 2012-2013. Availableat: http://www.parliament.uk/documents/joint-committees/communications-data/writtenevidence-Volume.pdf.

House of Lords Select Committee on the Constitution. 2009. Surveillance: Citizens and theState. February 6. Published by the Authority of the House of Lords. Available at: http://www.publications.parliament.uk/pa/ld200809/ldselect/ldconst/18/18.pdf.

Hunton & Williams LLP. 2014. A Risk-based Approach to Privacy: Improving Effectiveness in Practice. Centre for Information Policy Centre. 19 June 2014. Available at: https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/white_paper_1-a_risk_based_approach_to_privacy_improving_effectiveness_in_practice.pdf.

Hunton & Williams LLP. 2016. Risk, High Risk, Risk Assessments and Data Protection Impact Assessments under the GDPR. CIPL GDPR Interpretation and Implementation Project. 21 December 2016. Available at: https://www.huntonprivacyblog.com/wp-content/uploads/sites/18/2016/12/cipl_gdpr_project_risk_white_paper_21_december_2016.pdf.

Hutchinson, T. 2015. The Doctrinal Method: Incorporating Interdisciplinary Methods in Reforming the Law. Terry Hutchinson. Available at: https://www.elevenjournals.com/tijdschrift/ELR/2015/3/ELR-D-15-003_006.

Hutchinson, T. and Duncan, N. 2013. Defining and describing what we do: doctrinal legal research. *Deakin Law Review* 17(1): 83-119. ICO. 2011. Data sharing code of practice. Available at: http://tinyurl.com/ICO-SHARE.

Hyde, K.F. 2000. Recognising deductive processes in qualitative research. *Qualitative Market Research: An International Journal* 3 (2): 82-90.

ICO. 2012a. Anonymisation: Managing data protection risk code of practice. Available at: https://ico.org.uk/media/1061/anonymisation-code.pdf.

ICO. 2012b. Determining what is personal data. Version 1.1. Available at: http://tinyurl.com/ICO-WHATISPD.

ICO. 2012c. Guidance on data security breach management. Available at: http://tinyurl.com/ICO-BREACHES.

ICO. 2012d. Initial analysis of the European Commission's proposals for a revised data protection legislative framework. V1.0. 27 February 2012. Available at: https://wiki.laquadrature.net/images/1/12/Ico_initial_analysis_of_revised_eu_dp_legislative_proposals.pdf.

ICO. 2012e. Practical Guide to IT Security. April 2012. Available at: https://ico.org.uk/media/for-organisations/documents/1575/it_security_practical_guide.pdf.

ICO. 2013. Proposed new EU General Data Protection Regulation: Article-by-article analysis paper. 12 February 2013. Available at: https://ico.org.uk/media/about-the-ico/documents/1042564/ico-proposed-dp-regulation-analysis-paper-20130212.pdf.

ICO. 2014a. Big Data and Data Protection. 20140728, Version: 1.0. July 28. Available at: https://ico.org.uk/media/for-organisations/documents/1541/big-data-and-dataprotection.pdf.

ICO. 2014b. Conducting Privacy Impact Assessments Code of Practice. V1.0. February 2014. Available at: https://ico.org.uk/media/for-organisations/documents/1595/pia-code-of-practice.pdf.

ICO. 2014c. Data controllers and data processors: what the difference is and what the governance implications are. Available at: http://tinyurl.com/ICO-CONT-PROC.

ICO. 2016. Guide to Data Protection version 2.4. Available at http://tinyurl.com/ICO-DPG2-4.

InfoDev/ITU ICT Regulatory Toolkit. 2013. New Technologies and Their Impact on Regulation. Available at: http://www.ictregulationtoolkit.org/1.7.

Intelligence and Security Committee of Parliament. 2015. Privacy and Security: A modern and transparent legal framework. Presented to Parliament pursuant to Section 3 of the Justice and Security Act 2013. March 12.

International Working Group On Electronic Authentication. 1999. International Consensus Principles for Electronic Authentication. Internet Law & Policy Forum. April 23. Available at: http://www.ilpf.org/events/intlprin.htm.

IPCC. 2007. Regulations and Standards. IPCC Fourth Assessment Report: Climate Change2007 - Mitigation of Climate Change. Available at: https://www.ipcc.ch/publications_and_data/ar4/wg3/en/ch13s13-2-1-1.html.

ISP Review. 2010a. EC to Monitor DPI CView Trial on Virgin Media UK Broadband ISP Users. January 28. Available at: http://www.ispreview.co.uk/story/2010/01/26/ec-tomonitor-dpi-cview-trial-on-virgin-media-uk-broadband-isp-users.html.

ISP Review. 2010b. Virgin Media UK Halt Broadband ISP Trial of CView DPI to Track Illegal File Sharing. October 10. Available at: http://www.ispreview.co.uk/story/2010/10/01/virgin-media-uk-halt-broadband-isp-trial-ofcview-dpi-to-track-illegal-file-sharing.html.

Iversen, A., Liddell, K., Fear, N., Hotopf, M. and Wessely, S. 2006. Consent, confidentiality, and the data protection act. *British Medical Journal*, *332*(7534): 165-169, DOI: 10.1136/bmj.332.7534.165.

Jamieson, M. 2001. Liability for defective software. The Journal of the Law Society of Scotland. May 1. Available at: http://www.journalonline.co.uk/Magazine/46-5/1000702.aspx.

Japanese Information Security Policy Council. 2010. Information Security Strategy for Protecting the Nation. Available at: www.nisc.go.jp/eng/pdf/New_Strategy_English.pdf.

Jason, A. 2011. *The Basics of Information Security*. Syngress, Waltham, MA. Kahneman, D. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Jentzsch, N. 2016. Competition and Data Protection Policies in the Era of Big Data: Privacy Guarantees as Policy Tools. Brussels Privacy Symposium on Identifiability: Policy and Practical Solutions for Anonymisation and Pseudonymisation. November 8. Brussels, Belgium. Available at: https://fpf.org/wp-content/uploads/2016/11/Jentzsch_Ident_Workshop_Paper_2016_V8_FINAL-I.pdf.

Kang, D.H., Oh, J.T. and Kim, K.Y. 2005. Application Protocol-Based Anomaly Detectionfor High Speed Network. Information Security Research Division Electronics and Communications research Institute, Danjeon, Korea. Available at: http://www.apnoms.org/2005/technical/8_4.pdf.

Kannecke, U. and Körber, T. 2008. Technological Neutrality in the EC Regulatory Framework for Electronic Communciations: A Good Principle Widely Misunderstood. *European Common Law Review* 330-337.

Katsh, M.E. 1995. *Law in a digital world*. New York Oxford: Oxford University Press.

Kerikmäe, T. 2014. Regulating eTechnologies in the European Union: Normative Realities and Trends. Springer International Publishing. Available at: http://www.springer.com/gb/book/9783319081168.

Kerr, O.S. 2004. The Fourth Amendment and New Technologies: Constitutional Myths andthe Case for Caution. *MlCH. L. REV*. 102:801.

Kim, H., Claffy, K., Fomenkov, M., Barman, D., Faloutsos, M. and Lee K. 2008. Internet traffic classification demystified: myths, caveats, and the best practices. In Proceedings of the 2008 ACM CoNEXT conference. Madrid, Spain, December 09–12, 2008. CoNEXT ‘08. ACM, New York, NY, 1 –12.

Klein, J. and Newell, W. 1997. Advancing interdisciplinary studies. In J. Gaff and J. Ratcliffe (Eds.). *Handbook of the undergraduate curriculum: A comprehensive guide to purposes, structures, practices, and changes*. San Francisco: Jossey-Bass. pp. 393-415.

Klingbeil,M. 2010. SmartRegulation. Available at: https://www.oecd.org/regreform/policyconference/46528683.pdf.

Klitou, D. 2014. Privacy by Design and Privacy-Invading Technologies: Safeguarding Privacy, Liberty and Security in the 21st Century. *Legisprudence* 5(3): 297-329.

Koenig, C., Bartosch, A., Braun, J.D. and Romes, M. 2009. *EC Competition and Telecommunications Law* (2nd ed.). International Competition Law Series 6. The Netherlands: Kluwer Law International.

Koger, J.L. 2001. You Sign, E-SIGN, We all Fall Down: Why the United States Should NotCrown the Marketplace as Primary Legislator of Electronic Signatures. *Transnational Lawand Contemporary Problems* 11:419-516.

Koops, B.J. 2006. Should ICT Regulation be Technology-Neutral?. In B.J. Koops, M. Lips,C. Prins, and M. Schellekens (eds.). *Starting Points for ICT Regulation: deconstructing prevalent policy one-liners*. The Hague: TMC Asser. pp. 77-108.

Korf, D. 2002. Study on Implementation of Data Protection Directive – Comparative Summary of National Laws. September 2002. Available at: http://194.242.234.211/documents/10160/10704/Stato+di+attuazione+della+Direttiva+95-46-CE.

Krebs, D. 2013. Privacy by Design: Nice-to-Have or a Necessary Principle of Data ProtectionLaw?. *JIPITEC* 4. Available at: http://ssrn.com/abstract=2057937.

Kuhn, T. (3rd ed.). 1996. *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press.

Kumar, D. 1998. Problems With Code-Based Regulation. Available at: https://cyber.law.harvard.edu/fallsem98/final_papers/Kumar.html.

Kumar, S., Turner, J. and Williams, J. 2006. Advanced algorithms for fast and scalable deeppacket inspection. In Proceedings of the 2006 ACM/IEEE symposium on architecture fornetworking and communications systems. San Jose, California, USA, December 03–05, 2006. ANCS ’06, 2006.

Kuner, C. 1998. Remarks of the German Government on the EU Draft Directive concerningElectronic and Digital Signatures. April 8. Available at: http://www.kuner.com/data/sig/verbrauc.htm.

Kuner, C. 1999. German Consumer Association Denounces EU Draft Digital Signature Directive. Available at: http://www.kuner.com/data/sig/gov_ger_eu-draft.htm.

Lagos, Y. and Polonetsky, J. 2013. Public vs. Nonpublic Data: The Benefits of Administrative Control. 66 *STAN. L. REV. ONLINE* 103. Available at: https://www.stanfordlawreview.org/online/privacy-and-big-data-public-vs-nonpublic-data/.

Lah, F. 2008. Are IP Addresses Personally Identifiable Information?. *I/S: A Journal for Lawand Policy for the Information Society* 4(3): 681-707.

Lane, J., Stodden, V., Bender, S. and Nissenbaum, H. 2014. *Privacy, Big Data, and the Public Good.* Cambridge: Cambridge University Press.

Langheinrich, M. 2001. Privacy by Design - Principles of Privacy-Aware Ubiquitous Systems. Available at: http://www.vs.inf.ethz.ch/publ/papers/privacy-principles.pdf.

Leenes, R. 2011. Framing Techno-Regulation: An Exploration of State and Non-State Regulation by Technology. *Legisprudence* 5(2): 143-169.

Lessig, L. 1997. Tyranny in the Infrastructure: The CDA was Bad, but PICS May Be Worse.Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=11470.

Lessig, L. 1999. The Law of the Horse: What Cyberlaw Might Teach. *Harvard Law Review*113 (2): 501-549.

Lessig, L. and McChesney, R.W. 2006. No Tolls on the Internet. The Washington Post, June8. Available at: http://www.washingtonpost.com/wpdyn/content/article/2006/06/07/AR2006060702108.html.

Levin, O. and Salido, J. 2016. The Two Dimensions of Data Privacy Measures. Corporate, External and Legal Affairs, Microsoft. Brussels Privacy Symposium on Identifiability: Policy and Practical Solutions for Anonymisation and Pseudonymisation. November 8.

Brussels, Belgium. Available at: https://fpf.org/wp-content/uploads/2016/11/The-Two-Dimensions-of-Data-Privacy-Measures.pdf.

Lewis, J. A. March 2009. Innovation and Cybersecurity Regulation. Washington, DC: Centerfor Strategic and International Studies. Available at: http://csis.org/files/media/csis/pubs/090327_lewis_innovation_cybersecurity.pdf.

Lewis, J.A. and Baker, S. July 2013. The Economic Impact of Cybercrime and Cyber Espionage. Centre for Strategic and International Studies & McAfee. Available at: http://www.mcafee.com/us/resources/reports/rp-economic-impact-cybercrime.pdf.

Lie, E., Macmillan, R. and Keck, R. 2009. Cybersecurity: The Role and Responsibilities of an Effective Regulator. 9th ITU Global Symposium for Regulators. Beirut, Lebanon. November 2009. Available at : http://www.itu.int/ITU-D/treg/Events/Seminars/GSR/GSR09/doc/GSR-background-paper-on-cybersecurity-2009.pdf.

Lima, S. 2013. 3 Ways to Use DNS Rate Limit against DDoS Attacks. CloudShield, January29. Available at: http://www.cloudshield.com/blog/dns-security-expert-series/3-ways-touse-dns-rate-limit-against-ddos-attacks/.

Lipton, J.D. 2010. Digital Multi-Media and the Limits of Privacy Law. Case Western ReserveJournal of International Law 42(3): 551 -571.

Liska, A. 2003. The Practice of Network Security: Deployment Strategies for Production Environmnets. New Jersey: Pearson Education Inc.

Livesey, S. 2012. The EU's proposed new e-identification regime. *Out-Law.com*. Availableat: http://www.out-law.com/en/topics/tmt--sourcing/e-commerce/the-eus-proposed-new-eidentification-regime/.

Lovells, H. 2014. Technology neutrality in Internet, telecoms and data protection regulation.Global Media and Communications Quarterly. Available at: http://www.hoganlovells.com/files/Uploads/Documents/8%20Technology%20neutrality%20in%20Internet.pdf.

Lyon, D. 2008. *Surveillance Studies: An Overview*. Malden, MA: Polity Press.

Mackey, E. and Elliot, M. 2013. Understanding the Data Environment. *XRDS:Crossroads* 20 (1): 37-39.

Malin, B., Sweeney, L. and Newton, E. 2003. Trail re-identification: learning who you are from where you have been. Tech. Report No.LIDAP-WP12. Carnegie Mellon University, Laboratory for International Data Privacy. Pittsburgh, PA: March 2003.

Marshall, C. 2013. Don't Blame Open Recursives For DDoS Attacks, Why You Should Implement BCP38. *Dyn*, April 2. Available at: http://dyn.com/blog/ddos-attacks-bcp38-internet-security-cloudflare-downtime-managed-dns-open-recursives/.

Massiello, B. and Whitten, A. 2010. Engineering Privacy in a Age of Information Abundance. AAAI Spring Symposium: Intelligent Information Privacy Management. Available at:http://dblp.uni-trier.de/db/conf/aaaiss/aaaiss2010-5.html#MasielloW10.

Matzner, T., Masur, P.K., Ochs, C. and Von Pape, T. 2016. Do-It-Yourself Data Protection—Empowerment or Burden?. In S. Gutwirth, R. Leenes and P. De Hert (eds.). *Data Protection on the Move: Current Developments in ICT and Privacy/Data Protection.* Heidelberg Springer. pp. 277-306.

May, P.J. 2004. Performance-Based Regulation And Regulatory Regimes. 1 3th World Conference on Earthquake Engineering. Vancouver, B.C., Canada. August 1 -6. Paper No. 3254. Available at: http://www.iitk.ac.in/nicee/wcee/article/13_3254.pdf.

Mayer-Schönberger, V. and Cukier, K. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. London: Jonh Murray.

Maynard, A. 2011. Regulating emerging technologies – Science & Public Participation top anew White House set of principles. April 16. Available at: http://2020science.org/2011/04/16/regulating-emerging-technologies-science-publicparticipation-top-a-new-white-house-set-of-principles/.

McAskill, E. 2014. Vodafone feels Edward Snowden effect with surveillance revelations. *TheGuardian*, June 6. Available at:

http://www.theguardian.com/world/2014/jun/06/analysisvodafone-feels-edward-snowden-effects.

McCullagh, D. 2008. Q&A with Charter VP: Your web activity, logged and loaded. *CNET News*, May 15. Available at: http://news.cnet.com/8301 -13578_3-9945309-38.html.

McCullagh, K. 2007. Data sensitivity: proposals for resolving the conundrum. *Journal of International Commercial Law and Technology* 2 (4): 190-201. Available at: http://tinyurl.com/z874zjp.

McIntyre, T. J. 2013. Child Abuse images and Cleanfeeds: Assessing Internet Blocking Systems. In I. Brown (ed.). *Research Handbook on Governance of the Internet*. UK: Edward Elgar Publishing Ltd. pp.277-308.
McSaty, A. 2011. Profiling Phorm: an autopoietic approach to the audience-as-commodity.*Surveillance & Society* 8(3): 311-322.

Mochalski, K. and Schulze, H. 2009. Deep Packet Inspection: Technology, Applications & NetNeutrality. *IPOQUE.* Available at: http://www.ipoque.com/sites/default/files/mediafiles/documents/white-paper-deep-packetinspection.pdf.

Moon, C.S. and Kim, S.H. 2014. A Study on the Integrated Security System based Real-timeNetwork Packet Deep Inspection. *International Journal of Security and Its Applications* 8(1): 113-122.

Morton, A. and Sasse, M.A. 2012. Privacy Is A Process, Not A PET: A Theory For EffectivePrivacy Practice. In Proceedings of the 2012 Workshop on New Security Paradigms
(NSPW '12). ACM, New York, NY, USA, 87-104. DOI=10.1145/2413296.2413305. Available at: http://doi.acm.org/10.1145/2413296.2413305.

Moses, L.B. 2007. Recurring Dilemmas: The Law's Race to Keep Up with Technological Change. UNSW Law Research Paper No. 2007-21. Available at: http://ssrn.com/abstract=979861.

Mueller, M. 2010. *Networks and states: The global politics of Internet governance*. Cambridge, Mass.: The MIT Press.

Mueller, M. 2011. DPI Technology from the standpoint of Internet governance studies: An introduction. Available at: http://dpi.ischool.syr.edu/Technology_files/WhatisDPI-2.pdf.

Narayanan, A. and Felten, E.W. 2014. No Silver Bullet: De-identification Still Doesn't Work. Available at: http://randomwalker.info/ publications/no-silver-bullet-deidentification.pdf.

Narayanan, A. and Shmatikov, V. 2008. Robust De-Anonymization of Large Sparse Datasets. Available at: https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf.

Narayanan, A. and Shmatikov, V. 2009. De-anonymizing Social Networks. Published in the30th IEEE Symposium on Security and Privacy, 2009. Available at: http://www.cs.utexas.edu/~shmat/shmat_oak09.pdf.

Narayanan, A. and Shmatikov, V. 2010. Myths and Fallacies of Personally Identifiable Information. *COMM. ACM* 53: 24.

Narayanan, A., Huey, J. and Felten, E.W. 2016. A Precautionary Approach to Big Data Privacy. In Gutwirth, S., Leenes, R., & De Hert, P. (eds.) *Data Protection on the Move:* Heidelberg Springer. pp. 357-386.

Naughton, J. 2008. Wikipedia censorship highlights a lingering sting in the tail. *The Guardian*, December 14. Available at: http://www.theguardian.com/technology/2008/dec/14/wikipedia-censorship-scorpionsvirgin-killer.

Nelson, G. 2015. Practical Implications of Sharing Data: A Primer on Data Privacy, Anonymization, and De-Identification. Available at: https://pdfs.semanticscholar.org/8a09/1b5cc4d3f861c0080d7b3ddf51b717244e6c.pdf.

Nickson, N. 2009. Virgin Media to begin CView trials. *Techradar*, December 10, Availableat: http://www.techradar.com/news/internet/virgin-media-to-begin-cview-trials-657287#null.

Nissenbaum, H. 2004. Privacy as Contextual Integrity. *Washington Law Review* 79 (119): 101-139. Available at: http://tinyurl.com/j8xut58.

Nissenbaum. H. 2010. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Palo Alto, CA: Stanford University Press.

Norton, Q. 2013. The Dangers of Deep Packet Inspection. MaximumPC, February 5. Available at: http://www.maximumpc.com/article/columns/Deep_Packet_Inspection_2013.

O'Keefe, C. M. and Connolly, C. 2010. Privacy and the use of health data for research. *Med J Australia* 193: 537-541. Available at: http://tinyurl.com/zxgnhvq.

OECD. 2002. Guidelines for the Security of Information Systems and Networks: Towards aCulture of Security. Available at: http://www.oecd.org/sti/ieconomy/15582260.pdf.

OECD. 2011. Council Recommendation on Principles for Internet Policy Making. 13 December. Available at: http://www.oecd.org/internet/ieconomy/49258588.pdf.

OECD. 2012. Cybersecurity Policy Making at a Turning Point: Analysing a New Generationof National Cybersecurity Strategies for the Internet Economy. Available at: http://www.oecd.org/sti/ieconomy/cybersecurity%20policy%20making.pdf.

Office of the Privacy Commissioner of Canada. 2009. Review of the Internet Traffic Management Practices of Internet Service Providers. February 18. Available at: www.priv.gc.ca/information/research-recherche/sub/sub_crtc_090728_e.asp.

Office of the Privacy Commissioner of Canada. 2013a. Deep Packet Inspection Essay Project. Available at: http://www.priv.gc.ca/information/researchrecherche/dpi_intro_e.asp.

Office of the Privacy Commissioner of Canada. 2013b. What an IP Address Can Reveal About You. A report prepared by the Technology Analysis Branch of the Office of the Privacy Commissioner of Canada. May 2013. Available at: https://www.dal.ca/content/dam/dalhousie/doc/researchservices/Privacy%20Commission%20of%20Canada%20re%20IP%20Address%20(2013%20May).pdf.

Ohm, P. 2009. The Rise and Fall of Invasive ISP Surveillance. *University of Colorado LawLegal Studies Research Paper* No. 08-22.

Ohm,P. 2010. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review* 57: 1701-1777.

Oswald, M. 2014. Share And Share Alike? An Examination of Trust, Anonymisation and Data Sharing with Particular Reference to an Exploratory Research Project Investigating Attitudes to Sharing Personal Data with the Public Sector. *SCRIPTed* 11(3): 245-272. Available at: https://script-ed.org/wp-content/uploads/2014/12/oswald.pdf.

Out-Law.com. 2011. ISPs' traffic management may breach data protection and privacy laws,EU watchdog says. *Out-Law.com*, October 12. Available at: http://www.outlaw.com/en/articles/2011/october/isps-traffic-management-may-breach-data-protection-andprivacy-laws-eu-watchdog-says/.

Out-Law.com. 2012. Deep packet inspection standard cannot guarantee privacy, says academic. *Out-Law.com*, December 11. Available at: http://www.outlaw.com/articles/2012/december/deep-packet-inspection-standard-cannot-guaranteeprivacy-says-academic/.

Out-Law.com. 2014. Anonymous means NO identifying element left behind – EU handbook. *The Register*. February 4. Available at: http://www.theregister.co.uk/2014/02/04/new_data_protection_handbook_outlines_alternative_test_for_determining_anonymisation/.

Parsons, C. 2008. Deep packet inspection in perspective: tracing its lineage and surveillancepotentials. The New Transparency Surveillance and Social Sorting Working Paper,January10, Available at: http://www.christopherparsons.com/Academic/WP_Deep_Packet_Inspection_Parsons_Jan_2009.pdf.

Parsons, C. 2009. Deep Packet Inspection and Law Enforcement. July 2. Available at: http://www.christopher-parsons.com/deep-packet-inspection-and-law-enforcement/.

Parsons, C. 2012. Deep Packet Inspection and Its Predecessors. Available at: http://www.christopher-parsons.com/Main/wp-content/uploads/2013/02/DPI-and-ItsPredecessors-3.5.pdf.

Pearce, D., Campbell, E. and Harding, D. 1987. *Australian Law Schools: A Discipline Assessment for the Commonwealth Tertiary Education Commission*. Canberra : Australian Government Publishing Service.

Person, A.N. 2009. Behavioral Advertisement Regulation: How the Negative Perception ofDeep Packet Inspection Technology May Be Limiting the Online Experience. *Federal Communications Law Journal* 62(2), 435-464.

Pfitzmann, A. and Hansen, M. 2010. A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management. Version v 0.34. Available at: http://dud.inf.tu-dresden.de/Anon_Terminology.shtml.

Polonetsky, J., Tene, O. and Finch, K. 2016. Shades of Gray: Seeing the Full Spectrum of Practical Data Deidentification. 56 *SANTA CLARA L. REV*. 593.

Porter, T. 2010. The Perils of Deep Packet Inspection. *Symantec*, October 19. Available at: http://www.symantec.com/connect/articles/perils-deep-packet-inspection.

PRACTIS. 2011. Privacy Appraising Challenges to Technologies And Ethics. Final HorizonScanning Report. July 2011. Available at: http://practis.org/docs/PRACTIS%.

Privacy Protection Study Commission. 1977. Personal Privacy In An Information Society. July 1977. Available at: https://www.ncjrs.gov/pdffiles1/Digitization/49602NCJRS.pdf.

Radisys. 2010. DPI: Deep Packet Inspection Motivations, Technology, and Approaches forImproving Broadband Service Provider ROI. Radisys White Paper, September 2010. Available at: http://go.radisys.com/rs/radisys/images/paper-dpi-motivations.pdf. Rajab,A. 2009. Technological Neutrality. *Lex Electronica* 14(2). Available at: http://www.lex-electronica.org/docs/articles_236.pdf.

Ramos, A. 2009. Deep Packet Inspection Technologies. In H.F. Tipton & M. Krause (eds.) (6th ed.). *Information Security Management Handbook*. New York: Auerbach Publications.

Reed, C. 2007. Taking Sides on Technology Neutrality. *SCRIPT-ed* 4(3): 263-284. Availableat: http://www2.law.ed.ac.uk/ahrc/script-ed/vol4-3/reed.pdf.

Reed, D.P. 2008. Statement of Dr. David P. Reed to Subcommittee on Telecommunicationsand the Internet Committee on Energy and Commerce U.S. House of Representatives. July 17. Available at: http://energycommerce.house.gov/images/stories/Documents/Hearings/PDF/Testimony/TI/110-ti-hrg.071708.Reed%20-testimony.pdf.

Renals, P. and Jacoby, G. 2009. Blocking Skype through deep packet inspection. Paper presented at 42nd International Conference on System Sciences, Hawaii, 5–8 January. Available at: http://www.computer.org/csdl/proceedings/hicss/2009/3450/00/07-05-01-abs.html.

Repko,A.F. (2nd ed.). 2011. *Interdisciplinary Research: Process and Theory*. California: SAGE Publications, Inc.

Richards, C. 2012. Boyd's OODA Loop. Available at: http://www.jvminc.com/boydsrealooda_loop.pdf.

Richards, N. M. and King, J.H. 2013. Three Paradoxes of Big Data. *Stanford Law Review Online* 41. Available at: http://ssrn.com/abstract=2325537.

Riley, C.M. and Scott, B. 2009. *Deep packet inspection: The end of the Internet as we knowit?*. Florence, MA: Free Press.

Robertson, R.J. 1998. Electronic Commerce on the Internet and the Statute of Frauds. *S.C. L.REV.* 49:787.

Roman, J. 2013. Safeguarding ISPs from DDoS Attacks: The Role of Security Best Practices. Bank Info Security, May 21. Available at: http://www.bankinfosecurity.com/isp-securityneeds-to-be-improved-a-5773/op-1.

Room, S. 2014. The legal obligations for encryption of personal data in United States, Europe, Asia and Australia. 2014. Available at: http://www.xnetworks.es/contents/Vormetric/2014-The-legal-obligations-for-encryption-ofpersonal-data-in-Europe-Asia-and-Australia.pdf.

Rout, D. 2015. Developing a Common Understanding of Cyber security. *ISACA Journal* Vol. 6. Available at: https://www.isaca.org/Journal/archives/2015/volume-6/Pages/developing-a-common-understanding-of-cybersecurity.aspx.

RSA. 2014. To Regulate or Not to Regulate Cyber Security: That is the Question. RSA Conference, San Francisco. February 24-28. Available at: http://www.rsaconference.com/writable/presentations/file_upload/grc-w02-to-regulate-ornot-to-regulate-cyber-security-that-is-the-question.pdf.

Rubinstein, I. 2012. Regulating privacy by design. *Berkeley Technology Law Journal* 26:1409–1456.

Rubinstein, I. 2016. Framing the Discussion. Brussels Privacy Symposium on Identifiability: Policy and Practical Solutions for Anonymisation and Pseudonymisation. November 8. Brussels, Belgium. Available at: https://fpf.org/wp-content/uploads/2016/11/Rubinstein_framing-paper.pdf.

Rubinstein, I.S. 2013. Big Data: The End of Privacy Or A New Beginning?. *International Data Privacy Law*. January 25. Available at: http://idpl.oxfordjournals.org/content/early/2013/01/24/idpl.ips036.full.pdf+html.

Rubinstein, I.S. and Hartzog, W. 2015. Anonymization and Risk. New York University Public Law and Legal Theory Working Papers - Paper 530. Available at: http://lsr.nellco.org/cgi/viewcontent.cgi?article=1534&context=nyu_plltwp.

Russell, C., and W. Vaughan. 2003. *The Choice of Pollution Control Policy Instruments in Developing Countries: Arguments, Evidence and Suggestions*. International Yearbook of Environmental and Resource Economics, vol. VII. Cheltenham, UK: Edward Elgar.

Ryan, D.J. 2003. Two views on security software liability. Let the legal system decide. *IEEESecurity & Privacy* 99(1): 70-72.

Sabel, C.F. and Simon, W.H. 2011. Minimalism and Experimentalism in the Administrative State. 100 *GEO. L.J.* 53.

Sales, N.A. 2009. Run for the Border: Laptop Searches and the Fourth Amendment. *U. RICH. L. REV.* 43:1091.

Sales, N.A. 2013. Regulating Cyber-Security. *Northwestern University Law Review* 107(4):1503-1568.

Samuelson, P. 2000. Five Challenges for Regulating the Global Information Society. In C. Marsden (ed.). *Regulating the Global Information Society*. London: Routledge. pp. 317-332.

Savin, A. 2014. How Europe formulates Internet policy. *Internet Policy Review Journal onInternet Regulation* 3(1). February 26. Available at: http://policyreview.info/articles/analysis/how-europe-formulates-internet-policy.

Schneier, B. 2003. Liability changes everything. November. Available at: https://www.schneier.com/essays/archives/2003/11/liability_changes_ev.html.

Schneier, B. 2008. Software Makers Should Take Responsibility. July 17. Available at: https://www.schneier.com/essays/archives/2008/07/software_makers_shou.html.

Schouten, C. 2008. EU Failed to Apply Technology Neutrality in Regulating Communication Services. *Innovations Report*. November 12. Available at: http://www.innovationsreport.com//html/reports/communication-media/eu-failed-apply-technology-neutralityregulating-124187.html.

Schwartz, P. 2016. Risk and high risk: Walking the GDPR tightrope. May 29. Available at: https://iapp.org/news/a/risk-and-high-risk-walking-the-gdpr-tightrope/.

Schwartz, P. M. and Solove, D. J. 2011. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *New York University Law Review* 86: 1814-1894.

Science and Technology Committee of Parliament. 2015. Oral Evidence: Investigatory Powers Bill: Technology Issues. HC 573. November 10. Available at: http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/science-and-technology-committee/investigatory-powers-bill-technology-issues/oral/24378.pdf.

Shelanski, H. 2013. Information, Innovation, and Competition Policy for the Internet. *University of Pennsylvania Law Review* 161: 1663-1705.

Simitis, K. 1999. *Revisiting sensitive data*. Available at: http://tinyurl.com/j4w35bp.

Simon, W.H. 2015. The Organizational Premises of Administrative Law. *Law & Contemporary Problems* 78: 61-100.

Singleton, P. and Wadsworth, M. 2006. Confidentiality and consent in medical research: Consent for the use of personal medical data in research. *British Medical Journal* 333(7561): 255. Available at: http://tinyurl.com/jc8f3zo.

Smith, G. 2015. The Coming UK Surveillance Debate: Future-proofing. *Cyberleagle*. August 12. Available at: http://cyberleagle.blogspot.co.uk/2015/08/the-coming-uk-surveillancedebate_63.html.

Smith, G. 2016. The draft Investigatory Powers Bill - start all over again?. *Cyberleagle.* February 16. Available at: http://cyberleagle.blogspot.co.uk/2016/02/the-draftinvestigatory-powersbill.html?utm_source=feedburner&utm_medium=email&utm_campaign=Feed:+Cyberleagle+%28Cyberleagle%29.

Solove, D.J. 2002. Conceptualizing Privacy. *California Law Review* 90: 1087-1156.

Solove, D.J. 2004. *The Digital Person: Technology and Privacy in the Information Age*. New York: New York University Press.

Solove, D. J. 2008. *Understanding Privacy*. Cambridge, Mass.: Harvard University Press.

Spanish Government. 2011. Spanish Security Strategy: Everyone's responsibility. Available at: http://www.lamoncloa.gob.es/documents/estrategiaseguridad_baja_julio.pdf.

Stalder, F. 2002. The Failure of Privacy Enhancing Technologies (Pets) and the Voiding of Privacy. *Sociological Research Online* 7(2). Available at: http://www.socresonline.org.uk/7/2/stalder.html.

Stalla-Bourdillon, S. 2016. A call for a common techno-legal language to speak about anonymisation, pseudonymisation, de-identification… Could this be one of the biggest challenges brought about by the GDPR?. November 9. Available at: https://peepbeep.wordpress.com/2016/11/09/a-call-for-a-common-techno-legal-language-to-speak-about-anonymisation-pseudonymisation-de-identification-could-this-be-one-of-the-biggest-challenges-brought-about-by-the-gdpr/.

Stalla-Bourdillon, S. and Knight, A. 2016. Anonymous data v. Personal data–A false debate: An EU perspective on anonymisation, pseudonymisation and personal data. *Wis. Int'l LJ.* 2016.

Stalla-Bourdillon, S., Papadaki, E. and Chown, T. 2014. From Porn to Cybersecurity Passing by Copyright: How Mass Surveillance Technologies Are Gaining Legitimacy… The case of Deep Packet Inspection Technologies. *Computer Law & Security Review* 30 (2014): 670-686.

Stalla-Bourdillon, S., Papadaki, E. and Chown, T. 2015. Metadata, Traffic Data, Communications Data, Service Use Information... What is the Difference? Does the Difference Matter? An Interdisciplinary View from the UK. Serge Gutwirth & Ronald Leenes, Data Protection on the Move, Springer 2015. Available at SSRN: https://ssrn.com/abstract=2625181.

Statistics Canada. 2015. *What prevents a researcher from removing data from an RDC?*. Available at: http://www.statcan.gc.ca/eng/rdc/faq#a8.

Stevens, T. 2008. Phorm: a new dawn for web advertising?. *Computer Weekly*, April 22. Available at: http://connection.ebscohost.com/c/editorials/32043271/phorm-new-dawnweb-advertising.

Stewart, M.J. 2014. *Network Security, Firewalls and VPNs. Information Systems Security &Assurance Series.* London: Jones and Bartlett Learning LLC. Sweeney, L. 2000. Simple Demographics Often Identify People Uniquely. Carnegie Mellon University, Data Privacy Working Paper 3. Pittsburgh. Available at: http://dataprivacylab.org/projects/identifiability/paper1.pdf.

Stewart, R.B. 1981. Regulation, Innovation, and Administrative Law: A Conceptual Framework. 69 *Cal. L. Rev.* 1256.

Sumroy, R. and Cousin, R. 2016. Personal data, anonymisation and pseudonymisation under GDPR. July 2016. Available at: https://www.slaughterandmay.com/media/2535637/personal-data-anonymisation-and-pseudonymisation-under-the-gdpr.pdf.

Sweeney, L. 2002. k-Anonymity: A Model for Protecting Privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems* 10: 557.

Szongott, C., Henne, B. and Von Voigt, G. 2012. Big data privacy issues in public social media; In *6th IEEE International Conference on Digital Ecosystems Technologies (DEST)*, pp. 1-6. Campione d'Italia, Italy, 18 – 20 June 2012, IEEE.

Tanenbaum, A.S. and Wetherall, D.J 2010. *Computer Networks (5th ed.).* US: Pearson.

Tehan, R. 2013. Cybersecurity: Authoritative Reports and Resources. CRS Report for Congress, Prepared for Members and Committees of Congress. May 24. Available at: http://www.fas.org/sgp/crs/misc/R42507.pdf.

Tene, O. and Polotensky, J. 2012. To Track or "Do or "Do Not Track": Advancing Transparency and Individual Control in Online Behavioral Advertising. *Minnesota Journal of Law, Science & Technology* 13(1): 281-357.

Tene, O. and Polotensky, J. 2013. Big Data For All: Privacy and User Control In the Age of Big Data Analytics. *Northwestern Journal of Technology and Intellectual Property* 11(5): 239-273.

The Identity Projects. 2005. The Identity Project: An assessment of the UK Identity Cards Bill and its implications. Version 1.09, June. London: LSE Department of Information Systems.

The Identity Projects. 2007. Submission to the House of Lords Inquiry into the 'Impact of Surveillance and Data Collection'. London: London Scholl of Economics and Political Science Identity Project. Department of Information System.

Thordarson, J. 2015. 4 Challenges the new EU Data Protection Regulation will bring to IT Porfessionals. February 25. Available at: http://www.eudataprotectionregulation.com/#!4-Challenges-the-new-EU-Data-Protection-Regulation-will-bring-to-ITPorfessionals/cx28/04E3B410-8535-4443-84F7-B739F5A6E265.

Tockar, A. 2014. Riding with the Stars: Passenger Privacy in the NYC Taxicab Dataset. *NEUSTAR.* September 15. Available at: http://research.neustar.biz/author/atockar.

UK Cabinet Office. 2009. Cyber Security Strategy of the United Kingdom.Safety, Security and Resilience in Cyber Space. Available at: www.cabinetoffice.gov.uk/media/216620/css0906.pdf.

UK Cabinet Office. 2011. The UK Cyber Security Strategy. Protecting and promoting the UK in a digital world. Available at: www.cabinetoffice.gov.uk/resource-library/cyber-securitystrategy.

UK Government. 2013. Guiding principles on cyber security: Guidance for Internet Service Providers and Government. December 2013. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/265328/bis-13-1327-guiding-principles-for-cyber-security-isps-and-hmg-FINAL.pdf.

U.S. Department of Commerce. June 2011. Cybersecurity, Innovation and the Internet Economy.Internet Policy Task Force. Available at: http://www.nist.gov/itl/upload/Cybersecurity_Green-Paper_FinalVersion.pdf.

U.S. Department of Health, Education & Welfare. 1973. Records, Computers and the Rights of Citizens. Report of the Secretary's Advisory on Automate Personal Data Systems. June 25. Available at: https://www.justice.gov/opcl/docs/rec-com-rights.pdf.

U.S. White House. 2009. Cyberspace Policy Review: Assuring a Trusted and Resilient Information and Communications Infrastructure. Available at: www.whitehouse.gov/assets/documents/Cyberspace_Policy_Review_final.pdf.

U.S. White House. 2011. Fact Sheet: Cybersecurity Legislative Proposal. Available at: https://www.whitehouse.gov/the-press-office/2011/05/12/fact-sheet-cybersecuritylegislative-proposal.

U.S. White House. 2011. International Strategy for Cyberspace. Prosperity, Security and Openness in a Networked World. Available at: www.whitehouse.gov/sites/default/files/rss_viewer/international_strategy _for_cyberspace.pdf.

Upadhaya, G.R. 2012. BGP Best Practices for ISPs. Packet Clearing House. Available at: http://archive.apnic.net/meetings/22/docs/tut-routing-pres-bgp-bcp.pdf.

Vadhan, S., King, G., Sweeney, L., Airoldi, E. and Gasser, U. 2012. Project Description: Privacy for Social Science Research. National Science Foundation (NSF) Award No. 1237235. Available at: http://privacytools.seas.harvard.edu/files/privacytools/files/projectdescription_1.pdf?m=1363618082.

Van der Haar, I.M. 2007. Technological Neutrality; What Does it Entail?. TILEC Discussion Paper No. 2007-009. Available at: http://ssrn.com/abstract=985260.

Van Eijk, N.A. 2013. Duties of Care on the Internet. In J. Krüger, B. Nickolay and S. Gaycken (eds.). *The Secure Information Society: Ethical, Legal and Political Challenges*. London: Springer. pp. 57-82.

Van Eijk, N.A. and Van Engers, T.M. 2010. Moving Towards Balance: A Study into Duties of Care on the Internet. Institute for Information Law, Leibniz Centre for Law, University of Amsterdam. Available at: http://www.ivir.nl/publications/vaneijk/Moving_Towards_Balance.pdf.

Van Gestel, R. and Micklitz, H.W. 2011. Revitalizing Doctrinal Legal Research in Europe: What About Methodology?. *European University Institute Working Papers Law* 5:26.

Veraart, S. 2013. No cyber security without government imposed regulation. Workshop on Legal Frameworks and Cyber-crime. Available at: http://www.intgovforum.org/cms/wks2013/workshop_2013_status_list_view.php?xpsltipq_je=90.

Vixie, P. and Schryver, V. 2012. DNS Response Rate Limiting (DNS RRL). Available at: http://ss.vix.su/~vixie/isc-tn-2012-1.txt.

Vranaki, A., Heyder, M. and Bellamy, B. 2016. Implementing and Interpreting the GDPR: Challenges and Opportunities - Towards a Successful and Consistent Implementation of the GDPR. Centre for Information Policy Leadership Workshop Report. Amsterdam, Netherlands. March 16, 2016. Available at: https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_gdpr_project_amsterdam_workshop_report.pdf.

Warwo, A. 2012. What Is Deep Packet Inspection? *PCWorld*, February 1. Available at: http://www.pcworld.com/article/249137/what_is_deep_packet_inspection_.html.

Weisbrot, D. 2005. The Future for Institutional Law Reform. In B. Opeskin and D. Weisbrot (eds.). *The Promise of Law Reform*. Annandale: Federation Press. pp. 31-55.

Westervelt, R. 2013a. DHS Cybersecurity Official Says Industry Falling Behind Attackers. *CRN*. February 26. Available at: http://www.crn.com/news/security/240149341/dhscybersecurity-official-says-industry-falling-behindattackers.htm?itc=xbodyrobwes&itc=refresh.

Westervelt, R. 2013b. Presidential Cybersecurity Order: Channel Impact Depends On Implementation. *CRN*. February 26. Available at: http://www.crn.com/news/security/240149453/presidential-cybersecurity-order-channelimpact-depends-on-implementation.htm?itc=refresh.

Westervelt, R. 2013c. Security Pros Say Lack Of Skilled Workers Is The Biggest Threat. *CRN*. February 26. Available at: http://www.crn.com/news/security/240149303/securitypros-say-lack-of-skilled-workers-is-the-biggest-threat.htm.

Westervelt, R. 2014. Cybersecurity Balance: Too Much Regulation Will Stifle Economy, Stall Innovation. *CRN*. March 4. Available at: http://www.crn.com/news/security/300071938/cybersecurity-balance-too-much-regulationwill-stifle-economy-stall-innovation.htm.

White House Executive Office of the President. 2014. Big Data: Seizing Opportunities, Preserving Values. Available at: http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf.

Whitley, E.A. 2013. On technology neutral policies for e–identity: A critical reflection based on UK identity policy. *Journal of International Commercial Law and Technology* 8(2):134-147.

Williams, B.K. and Brown, E.D. 2014. Adaptive Management: From More Talk to Real Action. *Environ Manage.* 53(2): 465–479.

Williams, C. 2009a. Spook firm readies Virgin Media filesharing probes. *The Register*, December 7. Available at: http://www.theregister.co.uk/2009/12/07/detica_visit/.

Williams, C. 2009b. Virgin Media to trial filesharing monitoring system. *The Register*, November 29. Available at: http://www.theregister.co.uk/2009/11/26/virgin_media_detica/.

Willis, L.E. 2015. Performance-Based Consumer Law. *The University of Chicago Law Review* 82: 1309-1409.

Winterford, B., and Hill, J. 2008. ISP-level content filtering won't work. ZDNet Australia, October 30, 2008. Available at: http://www.zdnet.com.au/insight/communications/soa/ISPlevel-content-filtering- won-twork/0,139023754,339292158,00.htm.

World Economic Forum. 2013. Unlocking the Value of Personal Data: From Collection To Usage. Prepared in collaboration with The Boston Consulting Group. February 2013. Available at: http://www3.weforum.org/docs/WEF_IT_UnlockingValuePersonalData_CollectionUsage_Report_2013.pdf.

Wray, R. 2009. Phorm: UK faces court for failing to enforce EU privacy laws. *The Guardian*, April 14. Available at: http://www.theguardian.com/business/2009/apr/14/phorm-privacydata-protection-eu.

Wright, B. and Winn, J.K. (3rd ed). 1999. *The Law of Electronic Commerce*. Aspen Law and Business.

Wu, F. T. 2013. Defining Privacy and Utility in Data Sets. *U.COLO. L. REV.* 84: 1117-1152.

Yadron, D. 2014. Boards Race to Fortify Cybersecurity. *The Wall Street Journal*. June 19. Available at: http://www.wsj.com/articles/boards-race-to-bolster-cybersecurity-1404086146.

Yakowitz, J. 2011. Tragedy of the Data Commons. *HARV. J.L. & TECH*. 25. Available at: http://jolt.law.harvard.edu/articles/pdf/v25/25HarvJLTech1.pdf.

Zeadally, S. and Badra, M. 2015. *Privacy in a Digital, Networked World: Technologies, Implications and Solutions*. London: Springer.

Zittrain, J. 2008. *The Future of the Internet and How to Stop It.* New Haven & London: Yale University Press.