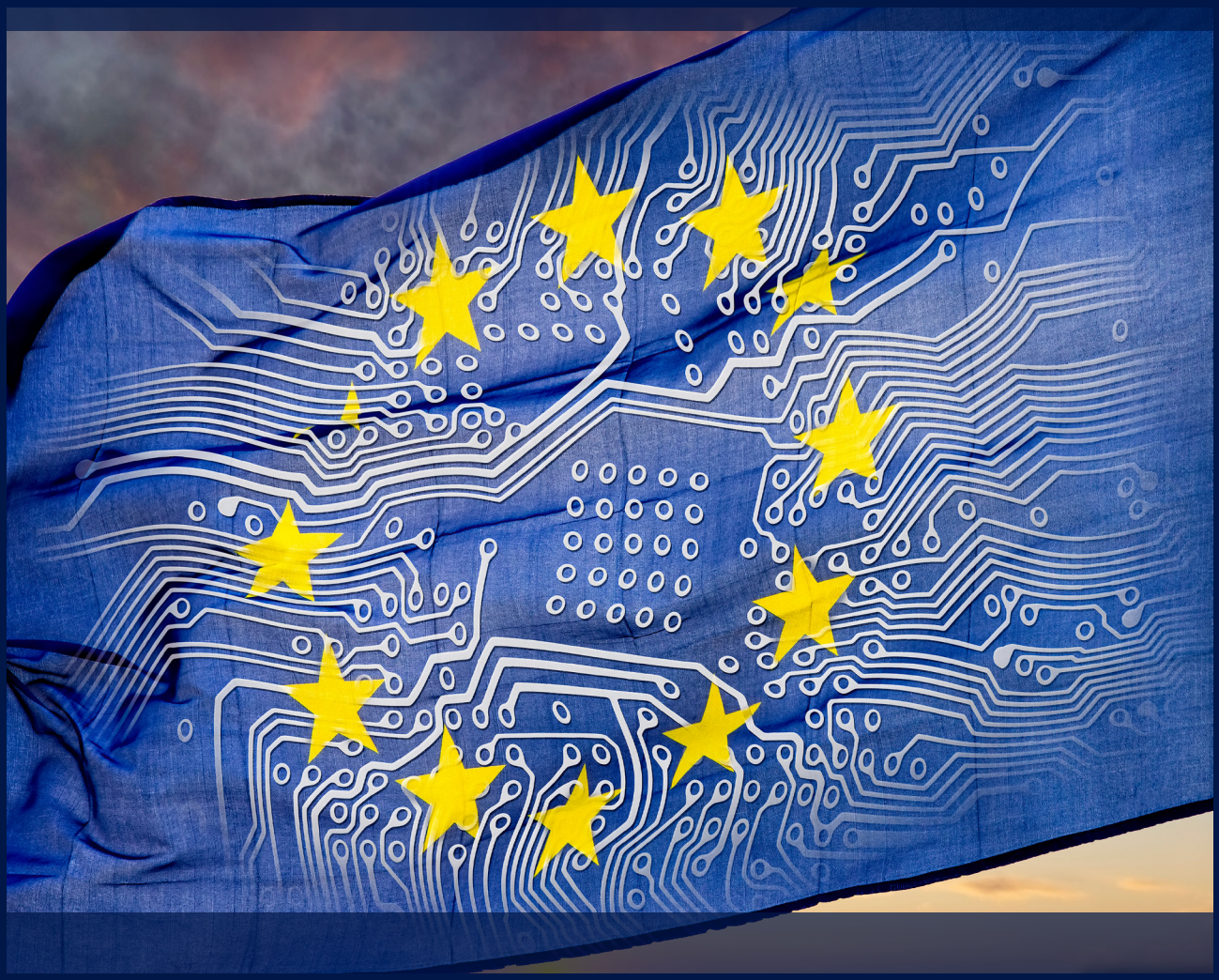
EUROPEAN DIGITAL INFRASTRUCTURE AND DATA SOVEREIGNTY

A POLICY PERSPECTIVE

FULL REPORT



CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

In the context of its activities in its strategic innovation area: Digital Industry, EIT Digital decided to launch a study focusing on the main policy challenges concerning the future of digital infrastructures also in relation to data governance and the consequences for Europe's sovereignty and economic strength. The study followed a scenario-based approach to structure and assess the potential impacts of policy measures with the main focus on various potential methods of regulation of modern digital infrastructures and the relation to personal data governance. Digital Enlightenment Forum was contracted to execute the study under the guidance of EIT Digital senior staff.

The very nature of the study strongly leads to strong attention to important European values, potential for global sovereignty and Europe's economic well-being. The study context is multi-dimensional and multi-disciplinary, and therefore demanded strong involvement by high-level experts from various disciplines for a profound assessment of the different dimensions and their interrelations. We were fortunate to have had such support throughout the study and in particular during through two Round Tables and electronic communication. We would like to acknowledge with gratitude the contributions of:

Michal Boni (former Member of European Parliament), Patrice Chazerand (Director Public Affairs, Digital Europe), Peter Dröll (EC/RESEARCH), Oliver Gray (Graywise, Consulting), Nina Hyvärinen (F-Secure), Volkmar Lotz (SAP), Evangelos Ouzounis (ENISA), Reinhard Posch (CIO Austrian Gov), Bart Preneel (Univ. Louvain BE), Luigi Rebuffi (EOS, ECSO).

Particular thanks go to Cristiano Codagnone (Univ. degli Studi di Milano), Thanassis Tiropanis (Univ. Southampton) and Yannis Stamatiou for their strong support providing a sound knowledge base and understanding of the complexity of the issues involved. And to Isaac Oluoch (Univ. of Twente, NL) for his help with analysing the literature. .

INTRODUCTION

The objective of this study report is to provide an analysis and set of scenarios as a tool for policy shapers on the future of digital infrastructures and personal data governance. Can we combine European value-based regulation with vibrant innovation for a sovereign, secure and trusted European society? In this introduction some of the main themes of the report are anticipated.

THE WORLD'S MOST VALUABLE RESOURCE IS NO LONGER OIL, BUT DATA...

The parallelism between oil and (personal) data was introduced in 2006 with the additional observation that, as with crude oil, data are valuable but if not refined cannot be used (Palmer, 2006). Data, and especially personal data, is the new crude asset for which a complex ecosystem of entities has emerged to collect, analyse and trade the value that may be extracted from it (often via behavioural insights). The importance of this view on data has been repeated by many afterwards.

The power of data is at the heart of what has been termed digital transformation and forms the pivotal element of the Fourth Industrial Revolution. Several trends have brought data to the core of innovation, including cheaper and more readily available storage and processing power, increasing availability of data through online social networks and Internet of Things (IoT), and data analytics improvements through the deep learning revolution. All of these are expected to be further powered by the deployment of 5G networks. These current and future trends have produced and are poised to further fuel network effects: data processing software becomes smarter as the volume of data contributed by people and their devices increases, which in turn makes the online services that employ such smart software more appealing and, consequently, attracts even more data.

DATA AND EMERGING DIGITAL INFRASTRUCTURES HOLD GREAT POTENTIAL FOR ECONOMY AND SOCIETY...

The data-driven digital transformation is a main source of innovation and according to recent experimental evidence has produced wide benefits for consumers that escape GDP measurement (Brynjolfsson et al., 2019).

▪ Modern Artificial Intelligence (AI) extracts value from data. More data means more accurate AI models, which in turn means potentially more benefits to society and business. We may expect a transition in production processes thanks to the adoption of AI, leading to significant economic growth and increase of economic output. The UK Government Office for Science expected 100 billion connected IoT devices already by 2020. ETNO, a major European telecoms association, expects 5G will underpin the rapid growth of the IoT (Palovirta & Grassia, 2019).

▪ The IoT will have a huge impact on automotive, industry, retail and smart building equipment.

▪ A document of the German Ministry of Economy officially presenting the project Gaia-X – a European cloud aiming to provide "the next generation of data infrastructure for Europe: a secure, federated system that meets the highest standards of digital sovereignty while promoting innovation" – underscores how cloud infrastructure can multiply the power and benefits as it encompasses the full spectrum of information technology and provides to firms in an efficient way (BMWi, 2019, p. 5).

▪ World-class connectivity in 5G is expected to be a key enabler of Europe's digital economy and of the increasing digitization of services and industrial processes, cutting costs and increasing efficiencies (See Palovirta & Grassia, 2019).

...YET DESPITE PROMISED BENEFITS, ANXIETY AND TENSIONS ARE RISING

The data economy is generating several major concerns and international tensions. First, much of the data powering those innovative services are personal or of a sensitive nature. The ongoing transformation affects all aspects of human reality (see Floridi, 2014 and Schwab, 2016, 2018), blurring the distinction between physical, digital and biological spheres. In this context, it comes as no surprise that the EU Ethics Advisory Group has expressed concern on the relationship between personhood and personal data, the risks of discrimination as a result of data processing, and the risks of undermining the foundations of democracy (Ethics Advisory

Group, 2018). In Europe, GDPR was a regulatory intervention to address some of those concerns. There has been support for the principles of GDPR to be adopted in the US, with Microsoft being one of the proponents of a US GDPR that lets people take control of their data. This follows the introduction of the California Consumer Privacy Act (CCPA), effective from 1 January 2020, which is seen as a first action taking some of the GDPR proposals on board.

Second, security concerns caused by increased cyber threats and attacks have become closely linked to the privacy and data protection issues, as has been noted by the European Commission (European Commission, 2017a, p. 11). As a result cybersecurity markets are growing rapidly: between 2016 and 2021 the global cybersecurity market is predicted to grow at a compound annual rate of 10% and be worth over $200 billion by 2021 (Morgan, 2015). While privacy and security concerns are tangible in everyday life, they are increasingly becoming entangled in geopolitical tensions. Ever since news spread in June 2013 on foreign surveillance (i.e., Wikileaks reports on US eavesdropping on Merkel etc.), the issue of technological sovereignty has emerged in Europe (Maurer et al., 2014) . Similar statements and proposals have continued on both sides of the Atlantic, some of which seem to be simply posturing for political consumption.

This is fuelling a debate on technological sovereignty and strategic autonomy. In the past three years strategic autonomy has acquired importance and has been mentioned by several world leaders, often in relation to digital technologies (Timmers, 2018; Timmers, 2019a, 2019b) . The German Economy Minister Peter Altmaier has been reported as arguing that a Europe-run cloud system could restore our digital sovereignty and counter unfair competition from state-controlled and state-subsidised companies from third countries (read China) and by market dominant online platforms (read US). On the same line of international political tensions, the Financial Times gave much emphasis to a letter allegedly written by Peter Altmaier to Margrethe Vestager (EU chief of the digital and competition dossier) arguing that "specific rules of behaviour need to be imposed on market-dominating online platforms" (Espinoza & Chazan, 2019). According to the two reporters, the letter contained statements such as "in light of current developments in the global data and digital economy, we require tougher oversight of abusive practices in order to maintain competition". According to a piece published in Politico, technological (incl. digital) sovereignty also figures high on the agenda of the new Commission and distrust of US tech giants in EU policy circles has allegedly "rekindled a protectionist zeal that many thought had been discarded long ago" (Scott, 2019). In the earlier cited piece, ETNO representatives urge the next European digital agenda to focus on IoT, 5G, and AI in order to close the gap with the US and China. Specular and equally adversarial posturing and actions can be found on the other side of the Atlantic. In the midst of this Transatlantic rift, came the news on 1 November, 2019 that a 'sovereign internet' law took effect in Russia: Moscow effectively gave itself the power to erect a sort of digital Iron Curtain around its networks, with the law allowing Roskomnadzor, Russia's telecoms agency, to shut the country off from external traffic exchange. In this context, it has been argued that several different 'Internets' are currently co-existing uneasily (O'Hara & Hall, 2018), including: the original libertarian Silicon Valley Open Internet, the Washington DC US Republican 'Commercial Internet', the 'Authoritarian Internet' epitomised by China, and what they call the 'Bourgeois Internet' envisaged by the European Commission ("where trolling and bad behaviour are minimised and privacy protected, possibly at the cost of innovation", ibid.), and now (our addition) the Iron Curtain Russian Internet.

Third, the oligopolistic access to valuable user data by few companies and their chosen partners has raised concerns on innovation and calls for regulation. The might of aforementioned network effects is evident in the impressive growth of giant technology companies (so called GAFAM: Google, Apple, Facebook, Amazon, Microsoft). The list of top 10 most valued public companies in the world based on market capitalisation, as of January 2019, is dominated by these and other US companies.

Fourth, there are issues of equitable access and persisting digital exclusion and divide. The digital transformation and revolution is not yet reaching everyone and the digital divide is taking new forms (United Nations, 2019, pp. 11-24). With specific regard to 5G, while a viable case for deployments of these new networks can be made for densely populated urban areas, there is a risk that rural and suburban areas get left out (ITU, 2018, p. 9).

Last but not least, the issues mentioned above threaten to weaken both nationally and internationally the public acceptability and trust that are key for the deployment and adoption of new emerging technological possibilities from which tangible benefit may accrue. Generalised and systemic trust, with the underpinning social capital, are the social glue that enables collaborative and productive practices in the digital ecosystem, since to a large extent this entails interaction among strangers with exchanges of very sensitive information.

In this report generalised and systemic trust is defined as an attitude entailing reliance on the benevolence of human nature or the attitude to give most people the benefit of the doubt.

Self-regulatory trust generating mechanisms (i.e., reputational ratings in online platforms) are usually extolled by industry and neo-liberal think tanks. Yet, such self-regulatory mechanisms of trust generation are vulnerable to abuse and not sufficient in the face of newly emerging possibilities with their related challenges in terms of data protection and cybersecurity, such as AI, IoT, 5G, etc. Trust must rest on more

solid pillars needing, in our view, to build a new governance framework (a mix of laws, regulation, awareness and education campaigns, as well as self-regulation and government-industry collaboration).

It is public actors that must lead this process to build the digital social capital to restore generalised and systemic trust both nationally and internationally.

EUROPEAN POSITIONING: TURNING DISADVANTAGES INTO ADVANTAGES?

The comparative disadvantages of Europe have been extensively discussed. Resorting to the initial metaphor of data being the oil of the 21st century, China and the United States are each large centralised markets, enabling the gathering of giant quantities of data to fuel their algorithms, whereas Europe is more fragmented, both in terms of markets and the dominant tech companies (O'Hara & Hall, 2018). As a consequence, the US and China are leading the 'refinery': the AI research and applications, as well as the specialised chips that run the latter. In the last 15 years the battle of domination in the digital landscape has led to the oligarchy of the American GAFAM, their Chinese counterparts (Alibaba, Baidu, Tencent, et al.) and other emerging, typically non-European, platforms. If the comparative disadvantages of Europe are well-known and documented, then the key question becomes: can Europe develop comparative advantages and exploit them?

One positive answer may come via 'creative parallelism' between the European regulatory frameworks for the protection of the physical landscape/environment (which is already quite developed) and of the digital landscape (which has only recently started to be developed). Europe played a leading role worldwide in the former, which has resulted in having a good competitive position globally in renewable energies and many 'vertical' applications compatible with this framework. Being first in developing and establishing this framework gave a considerable comparative advantage which was put to good and timely use by European players. As in former technology advances, current digital developments produce not only wonderful benefits and opportunities but also threats to personal liberties, democratic institutions and more.

This is where regulation of data protection and cybersecurity can come into play. The EU established as a first (and only) such initiative worldwide, the regulation for the protection of personal data (GDPR) which came into effect in May 2018. GDPR is joined by legislation for the use and reuse of public sector data (PSI: Public Sector Information), digital copyright, e-privacy and cyber security. Further steps on the regulatory framework for the use and re-use of privately-owned data are currently under discussion. It has been recognised that "the GDPR remains a source of advantage for the European Union — it is a leader in data protection because it is too large a market to ignore". It is being very influential worldwide, especially in Asia (Hogan Lovells, 2019) and has also stimulated debate and initiatives in the US (KPMG, 2018, pp. 8-9). The GDPR will also be a leading element in the positioning of Europe in the AI debate. A brief of the European Parliament on the EU ethical framework of AI, after describing its human-centric nature, adds the following statement: "While this approach will unfold in the context of the global race on AI, EU policy-makers have adopted a frame of analysis to differentiate the EU strategy on AI from the US strategy (developed mostly through private-sector initiatives and self-regulation) and the Chinese strategy (essentially government-led and characterised by strong coordination of private and public investment into AI technologies). In its approach, the EU seeks to remain faithful to its cultural preferences and its higher standard of protection against the social risks posed by AI – in particular those affecting privacy, data protection and discrimination rules – unlike other more lax jurisdictions" (European Parliament, 2019c).

From the above it becomes apparent that the EU and its Member States could jointly create a digital ecosystem that protects cherished values and at the same time helps jump-start innovation with built-in data protection and security ('by design and by default') which is trusted by citizens and entrepreneurs, thus providing new opportunities for European actors and platforms.

THE REGULATORS/INNOVATORS DILEMMA

It is typical for new technologies to be 'a Pandora's box' and entail risks and uncertainty. For early stage technologies regulators do not have answers to questions such as how will they further develop and be used? Who will benefit most? What could be possible harms? In addition to uncertainty on future developments, regulators may face a sort of technical impossibility in front of complex issues such as defining legally what it means to require algorithms to be transparent. Finally, current political and geopolitical tensions fuel rhetorical posturing rather than concrete actions. As a result, regulatory uncertainty hampers true innovators, favouring the status quo and its incumbents. This is the situation depicted in Figure 1, and note that in this context a 'laissez-faire' or 'leave it to the market' approach would not foster innovation for it would leave uncertainty and anti-competitive position intact. We will come back to this in Chapter 3.

The above themes will be further analysed in this report. To do so we give in Chapter 1 objective, scope (i.e. the dimensions of analysis), key concepts, and the approach. In pursuit of the objective, a methodology has been adopted that combines an analysis of a large number of written secondary sources with direct interaction with experts. A very comprehensive scoping review was performed that brings together the scientific literature with a vast array of other sources (industry reports, policy documents, etc. often called 'grey literature'). The results of the scoping are presented in Chapter 2

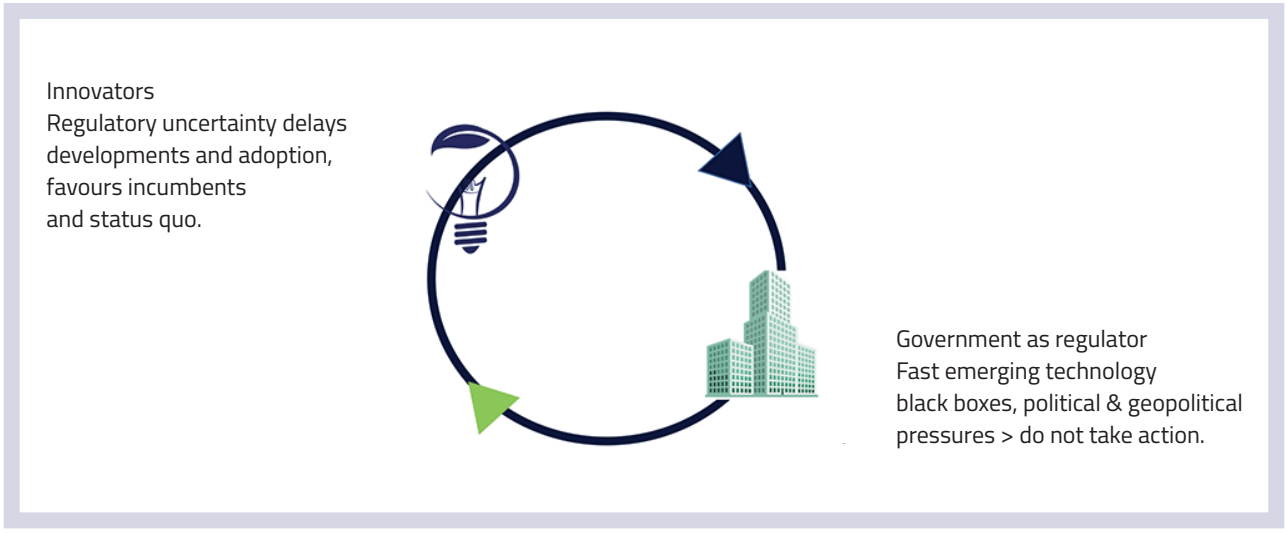


Figure 1: The regulators/innovators dilemma, source: adapted from (Deloitte, 2017, p. 2)

(and Annexes), leading to key trends and issues in the domain of digital infrastructure and data protection. Based on this analysis and the interaction with a number of experts from industry and academia, policy scenarios were developed using two dimensions of interventions: data governance and broadly defined digital infrastructure. These scenarios then lead to reflections on implications relevant for policy makers.

1. OBJECTIVE, SCOPE AND KEY CONCEPTS

1.1 Objective and scope

The objective of this study report is to provide an analysis and set of scenarios as a tool for policy shapers on the future of digital infrastructures and personal data governance. Can we combine European value-based regulation with vibrant innovation for a sovereign, secure and trusted European society?

The topic is complex. It spans from personal data governance to networks: mobile and fixed communication (spectrum, coverage, roll-out of 5G), Internet (net neutrality, domain name systems), data storage and management systems, cloud computing and data centres, applications, artificial intelligence (AI), Internet of Things (IoT), cybersecurity, and platforms. We need to delimit the scope and will focus on personal data governance (protection, sovereignty, security) and broadly defined digital infrastructures (selectively on 5G, IoT, Clouds, AI, and platforms, with cybersecurity considered horizontally).

GDPR and the debate that it has spurred in Europe and beyond justifies special attention to data governance in this report. Also KPMG placed data privacy among the ten regulatory challenges of 2019 (KPMG, 2018, pp. 8-9). Moreover, data-driven innovation is generally seen as one of the key engines of economic development, justifying to study how regulation of data protection can shape the use of data as raw material for the extraction of surplus and the role of dominant platforms as part of the digital infrastructure.

However, we need to note that in this report we will not address 'machine data' as it would broaden the scope beyond objective and capacities. We appreciate the importance for European Industry 4.0 of the trove of machine data (including sensor data) which is expected to increase exponentially with full development of the IoT. The EU Regulation on the free flow of non-personal data (Regulation (EU) 2018/1807, referred to as FFD Regulation) is applicable as of 28 May 2019. The FFD Regulation aims, among other things, to remove barriers to the free flow of non-personal data to foster the data economy, as was stressed in the 2017 Communication on Building a European Data Economy. Later the Commission issued a guidance to explain how the FDD regulation and GDPR interact. The guidance defines non-personal data as follows:

a) Data which originally did not relate to an identified or identifiable natural person, such as data on weather conditions generated by sensors installed on wind turbines, or data on maintenance needs for industrial machines;

b) Data which was initially personal data, but later made anonymous.

Further, it discussed the concept of mixed datasets, for in many cases datasets may comprise both personal and machine data. Mixed datasets will become increasingly common as a result of developments of AI and big data analytics. For example, the data recorded by a car including technical machine data (such as the tyre pressure), together with behavioural personal data (where and when the driver was at a certain place). This can also include a company's knowledge of IT problems and solutions based on individual incident reports. Currently, the interaction between the GDPR and FDD does not require separate processing of personal data and machine data. There is no obligation to split datasets, but the situation remains not fully defined for mixed datasets: the FFD Regulation applies to the non-personal data part of the dataset; the GDPR's free flow provision applies to the personal data part of the dataset. Yet, if the machine data parts and the personal data parts are inextricably linked, the data protection rights and obligations stemming from the GDPR fully apply to the entire mixed dataset, including in cases where personal data represents only a small part of the whole. Obviously, mixed datasets remain a thorny issue.

Nevertheless, the main analysis in this report considers only personal data governance. Although we will shortly come back to the issue of machine data in Section 3.2 when we discuss the implications of the scenarios presented.

As the main dimension of this analysis, digital infrastructures require careful definition. A simple definition used by economists to refer to infrastructure in general is "longer lived capital intensive systems and facilities" (Stupak, 2015, p. 1). According to Fourie (2006) infrastructure is defined by two dimensions: 'capitalness' and 'publicness' (referring to the social significance of the infrastructure and not necessarily to ownership). Applying these two dimensions, Palei (2015, p. 169) proposed the 2*2 typology reported below, where broadly-defined telecommunications figure among the high/high type.

Capitalness / Publicity	HIGH LEVEL	LOW LEVEL
HIGH LEVEL	Roads, highways, railways, airports, ports, electricity, water and sewerage, telecommunications	Schools, hospitals, parks, courts, museums theatres, libaries, universities, hospitals
LOW LEVEL	Industrial infrastructure	Fountains and statues

Table 1: Typology of infrastructures, source: adapted from Palei (2015, p. 169)

However, defining digital infrastructure raises some peculiar characteristics of the 'digital' aspect and moreover, the term 'digital infrastructure' is inconsistent and differs in current use. Digital 'things' are often of great public and social significance but relatively low in Capitalness (i.e. in terms of initial investment and maintenance costs, and not in terms of market capitalisation). For example: digital platforms or cloud computing that are marketed as a cheap solution for firms but are in aggregate strategic digital infrastructures. Following the above typology only very capital-intensive networks (fibre, 5G, etc.) would qualify as digital infrastructure. Nevertheless, intangible constructs, such as digital structural reform and the Digital Single Market, have very sizeable economic impact (Lorenzani & Varga, 2014). Hence, considering as digital infrastructure only the most tangible and physical aspects would be reductionist.

Concerning terminological practices, there are countless and inconsistent usages and definitions of the expression 'digital infrastructure', from very simple ones available on the web, such as "the ability to store and exchange data through a centralised communication system" or simple list of items (i.e., Internet backbone, fixed broadband, mobile, satellite, data centres, cloud computing, platforms, systems, apps, API & integration, user devices, IoT), to more formal but still debatable definitions. We take a pragmatic approach and refer in plural to 'digital infrastructures'. Going from the more to the less tangible, in this report digital infrastructures can include: fixed and mobile networks, IoT, applications and platforms, artificial intelligence. In Chapter 2 we will mainly focus on 5G, clouds and data centres, and algorithms and machine learning for their data protection implications. Regardless of their level of 'capitalness', all of these elements are characterised by four common characteristics: a) they have all been referred to as 'digital infrastructure' at least in more than two of the reviewed sources; b) they are all mentioned and debated in discussion of technological sovereignty, strategic autonomy, and cybersecurity; c) they all have great public and social significance, be it as enabler of economic and social activities relevant for all economic and societal actors or as a potential source of risks in need of some regulatory intervention; d) they are all expected to greatly contribute to economic growth and competitiveness.

The importance of digital infrastructures is clearly related to their potential societal and economic impact. There is a vast body of literature on the economic impact of infrastructure in general that is beyond our scope to review in depth here. There is broad consensus among economists that infrastructure has a clear impact on economic growth and competitiveness (Palei, 2015). There is ample evidence that infrastructures have a multiplier effect on economic output both in the short and long term, although the empirics on the size of such effect is not consolidated (Stupak, 2018). The various mechanisms through which the potential economic impacts of digital infrastructures unfold are the following: a) they expand capacity by increasing the efficiency of other existing infrastructures and lead to the emergence of new ones; b) they may save time, increase convenience, simplify operations, and can lead to more informed decisions; c) they save costs by decreasing waste and increasing efficiency allowing for more flexibility in the provision of goods and services; d) they could (i.e. provided that security is ensured) improve reliability, reducing volatility and uncertainty. Digital infrastructures are considered as the key driver of competitiveness, since it is the central and connecting infrastructure that enables gains in most other areas as depicted below.

1.2 Makers, shapers, and users

Figure 3 depicts the main stakeholders in the Digital Infrastructures domain (World Economic Forum, 2014). It is a starting point on which we flesh out our typology of players as including makers, shapers, and users.

Makers include technology innovators and solution providers: technology developers, communication services providers (CSPs), digital services and content providers, as well as hardware and software manufacturers (infrastructure devices and equipment, system software and support component manufacturing). Each digital infrastructure presents its own specificity and hence makers. For instance, the 5G value chain is particularly extended, encompassing mobile network operators, suppliers of mobile network operators, manufac-

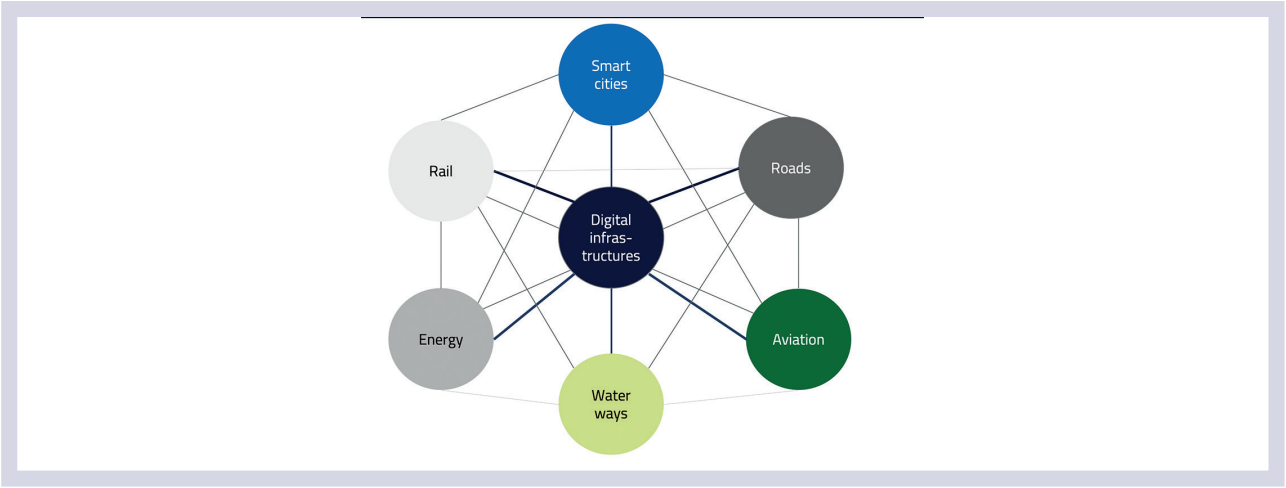


Figure 2: Digitization as key connecting infrastructure, source: adapted from Rudas et al. (2019)

turers of connected devices, service and content providers, and end-users. As we shall see later, this extended value chain raises security concerns, including security makers. The value chain and array of stakeholders, including the makers, for IoT is complex and fragmented, so far preventing the emergence of common standards for interpreting the data from devices or for connecting them to one another (Walport, 2014, p. 16). Competing platforms and industrial coalitions are emerging; and there are many different infrastructures that will form IoT networks.

Key makers include the dominant platforms and tech giants already having access to large amounts of data. With respect to data though, GAFAM is not a unified block (Arrieta-Ibarra et al., 2018; Faravelon et al., 2016). There is a clear difference both in business model and in level of access to data (and advancement on machine learning) between Google and Facebook on the one hand, and Apple and Microsoft on the other, with Amazon somewhere in the middle. Moving, for instance, to putting a price on users' data ('data as labour' model as proposed in Arrieta-Ibarra et al., 2018) would penalise Google and Facebook but may enable the other three giants to compete better in the data economy (Arrieta-Ibarra et al., 2018). Apple and Amazon are not major web actors, Google oversteps the influence of other platforms: Facebook only has half of Google's influence and Amazon does not rank in the Top 25 sites in terms of traffic (Faravelon et al., 2016, p. 27). So, access to data is differentiated and one may expect that their position with respect to regulation would not be monolithic as earlier cited support by Microsoft for a US GDPR shows.

AI start-ups are also makers, though constrained by limited access to data to train and improve their machine learning systems. Moving to makers that are more specific for cybersecurity the following distinction has been proposed: cybersecurity firms, internet technology firms, and internet-adjacent firms. Cybersecurity firms work for commercial or government clients providing products and/or services (i.e., Darktrace, FireEye, Palantir, Qadium, and Kaspersky Lab, Symantec, F-Secure, etc.). Internet technology firms are those mostly involved in the 'big data' space, including some of the tech giants. They may buy from cybersecurity firms or produce their own solutions. 'Internet-adjacent'

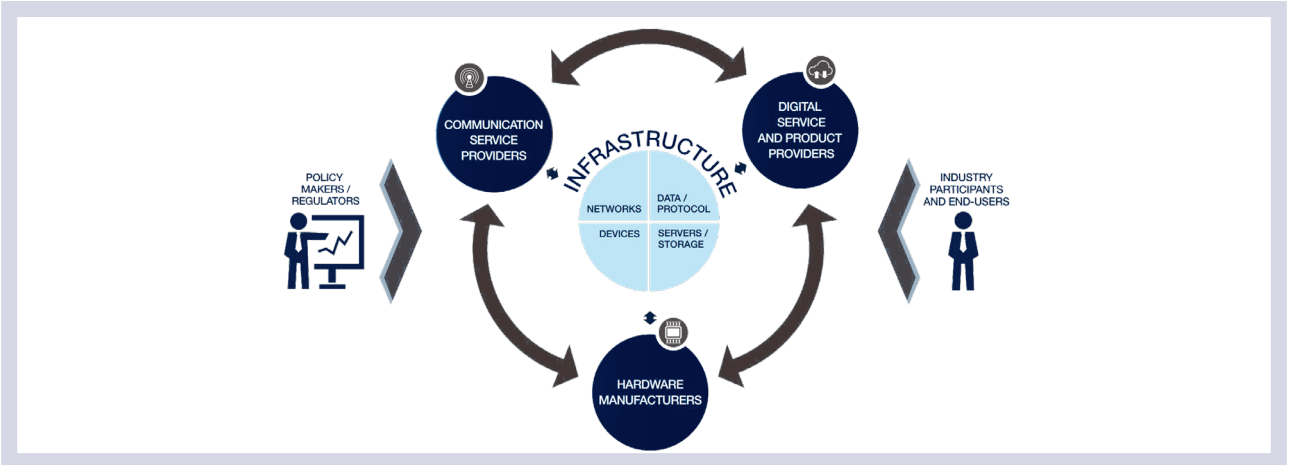


Figure 3: Digital ecosystem main stakeholders, source: World Economic Forum (2014)

firms – those that have digital components but have the core business outside the technology sector – could also be considered users (see below). Important examples are also large European companies like Thales or Siemens, working on large technical security and eID products. One could argue that this category includes all firms that are not makers of technology, since digital presence and activities are now widespread. Possibly a particular case is that of firms that will rely increasingly on IoT, such as energy companies: these may have special cybersecurity needs and may in the future produce their own solutions.

Shapers play a major role with respect to digital infrastructure, mainly in defining the specifications and regulating them. These include governments and other public sector actors. Governments usually have three principal roles: policy making, regulating, and owning digital infrastructure components and services (e.g. eGovernment). Other non-governmental agencies also have a crucial role as shapers. These include industry associations, standardisation bodies, and multiple stakeholder associations such as the World Wide Web Consortium (W3C), the Internet Corporation for Assigned Names and Numbers (ICANN) and the International Telecommunication Union (ITU). Large (e.g. US and Chinese) tech companies currently have access to and control of most of the data and enact their own policies in various respects (i.e., identity management, cybersecurity). They also play a de facto shaper role in addition to their maker role. They moved before regulation and created what are by now infrastructural solutions that may be hard to undo. Their lobbying power justifies treating them also as shapers.

Governmental shapers allegedly seek to protect the interests of the users/citizens and of economy and society as a whole when market failures (information asymmetry, externality, market power and imperfection) emerge. They may create conditions for equal access to innovation opportunities, thus addressing the interests of at least some groups of makers. Governments are also interested in both data protection and security as a matter of preserving the function and the values of democracy. It must be stressed that, on the issue of protecting the interests of the users, governmental shapers are seriously challenged by the capacity of tech giants to use rhetorical framing as a powerful lobbying strategy (Codagnone, 2017; Codagnone et al., 2018). This is particularly evident in the instrumental use of the rhetoric of the open Internet and of the interest of users receiving free services, where the original libertarian ethos of Silicon Valley is enlisted to defend powerful commercial interests (Zuboff, 2019).

Users include a wide range of players: from private individuals, ICT using firms, as well as governmental agencies. In Annex (section 4.1) we present detailed statistics and insights from the economics of privacy and information and from behavioural economics that shed light on attitudes and actual behaviours with respect to personal data and security, which we summarise here.

In general, it can be assumed that private individuals are interested in having lightweight and easy to use online services while preserving privacy and security. Firms may want to be compliant with GDPR to avoid fines, as well as avoiding security breaches both to protect their assets and intellectual property and to avoid the direct and indirect costs of such breaches (i.e., especially reputational costs that, for listed companies may translate into sizeable monetary losses). On the other hand, government may be particularly interested as users of cybersecurity to protect critical infrastructures. Behavioural biases and the complexity of the issues involved make users vulnerable to information asymmetry, from which the economy and society as a whole may suffer negative externality justifying regulatory interventions. It is very difficult for individuals to make reasoned and rational decisions on collection and use of their personal data and the related expression of consent due to behavioural bias and the complexity of the issues at stake. Data security and privacy are largely credence characteristics even with direct partners in a transaction (see more in Section 4.1). A credence good is something intangible that consumers can hardly verify, e.g. eco-friendliness of a washing machine (ecolabels have been developed to obviate this problem). In addition, data persistence means that consumer valuation of an information flow is a function of the network of entities that access and use that flow. The complexity of this network, combined with the difficulty in credibly conveying and committing to these policies, creates an information problem.

This applies also for executives of firms deciding on an investment for protecting the personal data of the customers or for increasing cybersecurity as it would be difficult for them to fully consider all the costs and benefits of making or not making the investment. Lack of information, the steadily changing costs, and the same behavioural biases seen earlier may hinder a complete appraisal of costs and benefit to determine the ROI. Then procrastination and responsibility dumping may set in and the investment would be avoided. Finally, because the indirect intangible costs of a security breach (related to loss of reputation potentially reverberating into losses in stock markets) are possibly higher than the direct costs, firms falling victim of a security breach may avoid publicly reporting it. But security breaches do not only generate costs at firms which are directly affected. Interdependence between information systems allows breaches to propagate and negatively affect others (Kunreuther & Heal, 2003). In the language of public economics, a lack of firms' information security (for any of the reasons above, and either at the level of adoption of cybersecurity measures or at that of publicly reporting breaches) causes negative externalities in an economy. The presence of negative externalities justifies government intervention, for instance, in the form of laws aimed at reducing the costs of insecurity to society (Hiller & Russell, 2013).

1.3 Three ideal-typical models of regulation

In the current digital geopolitical arms race three main ideal-typical models of approaching the regulation and governance of the digital landscape can be identified as represented by Europe, USA, and China respectively. These models are further supported in the Annex by a section providing a regulatory landscape overview. Other countries may be placed in one of these ideal types, but such granular positioning is beyond our scope (some other Asian countries in addition to China are discussed in the Annex). In presenting these three models we take inspirations from O'Hara & Hall (2018) with, however, two differences. First, we do not tout the European model as 'bourgeoise' but rather as descending from, and embodying, the main principles and values of the 'European project'. Second, we do not distinguish between a libertarian Silicon Valley Open Internet and a commercial Washington D.C. model. As Zuboff has shown at length (2019), openness and libertarian values are still held by minority players in Silicon Valley whereas for the big companies they remain simply rhetorical instruments. There is now, apart from nuances, full convergence between Silicon Valley and Washington.

The **European approach** is value- and human rights-based and focusses on ethics and privacy, whereas the American tradition leans more toward liberty (Whitman, 2004). The GDPR has enshrined into EU law a universalistic approach to the protection of privacy, extending protection of its citizens in other jurisdictions and enlarging the right to be forgotten by moving the emphasis from de-listing to erasure (Politou et al., 2018). As stated by the Ethics Advisory Group "This new data protection ecosystem stems from the strong roots of another kind of ecosystem: the European project itself, that of unifying the values drawn from a shared historical experience with a process of industrial, political, economic and social integration of States, in order to sustain peace, collaboration, social welfare and economic development" (EDPS Ethics Advisory Group 2018, 6). The GDPR covers all data processing activities to anticipate and minimise risk. In recent years the EU competition approach has been more proactive and the Commissioner Margrethe Vestager has extended the Commission's anti-trust work against dominant firms, based on Article 102 of the EU Treaty, to pursue American tech giants on the grounds that they might swallow rivals or force them out of business, leaving consumers with a poorer standard of service (The Economist, 2017a).

In general, aside from the inspiring values, an observer may also interpret the evolution of the European model (also in other areas as cybersecurity and attempted taxation of digital services) as having an implicit geopolitical agenda. In this respect, it has been noted that one of the drivers of the EU cybersecurity policy has also been for Europe to be a stronger global actor in international diplomacy, development cooperation, defence and trade (Timmers, 2018, p. 364). This has brought the EU at times on a collision course with its American Atlantic partners as can be seen, for instance, in the attack by Washington think tanks funded by Google and other tech companies, such as the International Technology and Innovation Foundation (ITIF).

The **US model** is a mix of a technology- and commercially-driven approach. With respect to privacy the dominant view is to treat it as tort, where the victim must prove the harm, which is in line with the Silicon Valley attitude to disrupt and move fast before regulation intervenes (Zuboff, 2019). In this respect the approach is commercial and there is convergence of views between Silicon Valley and Washington, despite current political misalignment between the Republicans dominating Washington and the Democrats which are more favoured in Silicon Valley (O'Hara & Hall, 2018).

As illustrated in the regulatory landscape presented in Annex, one characteristic of the US model is the lack of one unified federal framework on issues such as data protection and cyber security and the presence of several state laws and other sources of regulation or self-regulation and standardisation. On the other hand, the earlier cited report by KPMG shows that as a result of Europe introducing GDPR and other measures, there is mounting pressure in the US for a federal standardisation on Data Privacy (KPMG, 2018, pp. 8-9) and Cybersecurity (KPMG, 2018, pp. 14-15).

The **Chinese model** promotes its own tech giants (Baidu, Tencent and Alibaba) which work under close governmental control. These companies are hungrier, less complacent, more vigorous, more eager for competition, and less constrained by mission statements and core values than their US or European counterparts (Lee, 2018). Data protection in China is not up to European standards in terms of values and rights. China's cybersecurity market is, to all intents and purposes, driven by government prerogatives (Cheung, 2018). It is dominated by large monopolies with links to the national security apparatus. As a consequence, there are few firms in the Chinese cybersecurity marketplace. This, some have argued, decreases competition with negative effects upon the provision of cybersecurity (Cheung 2018).

One advantage of China, besides the fact that tech giants are less constrained by regulation provided they do as told by the government, resides in implementation capacity. The period of massive AI breakthroughs is being superseded by an age of implementation, of applying and adapting the algorithms to the dull problems of everyday life (Lee, 2018). Here, China has the advantage in terms of both the national skillset and the numbers of scientists it can deploy (Lee, 2018). Another advantage for China is the vast amount of data, as its Internet economy generates far more data than any other. Lastly, unhindered by data protection regulation or noticeable public demand for privacy, data is gathered from many other sources, including closed circuit television.

2. PERSPECTIVES AND TRENDS

2.1 Digital infrastructures

2.1.1 5G

The deployment of 5G networks is of great policy significance both for their potential benefits and for the challenges they raise. This includes not only global competition but also security concerns and because networks play an increasing role in the soft and hard power relations between states and regions (Albrycht & Swiatkowska, 2019; ITU, 2018, 2019; NIS Cooperation Group, 2019).

In the European Commission Recommendation Cybersecurity of 5G network [(EU) 209/534 of 26 March 2019] 5G networks are defined as "*a set of all relevant network infrastructure elements for mobile and wireless communications technology used for connectivity and value-added services with advanced performance characteristics such as very high data rates and capacity, low latency communications, ultra-high reliability, or supporting a high number of connected devices. These may include legacy networks elements based on previous generations of mobile and wireless communications technology such as 4G or 3G. 5G networks should be understood to include all relevant parts of the network*".

5G offers policy-makers the possibility to empower citizens and businesses, by transforming their cities into smart cities and allowing them to share in benefits delivered by a data driven digital economy, while at the same time providing wireless operators the opportunity to move beyond providing connectivity services and develop new solutions and services for consumers through wired and wireless converged networks with integrated network management systems (ITU, 2018, p. 3). 5G networks will provide virtually ubiquitous, ultra-high bandwidth and low latency connectivity not only to individual users but also to connected objects.

Thanks to these technical characteristics, 5G networks in combination with the IoT are expected to serve a wide range of applications and sectors, including a range of services such as energy, transport, banking and health, as well as industrial control systems. The organisation of democratic processes, such as elections, is also expected to rely more and more on digital infrastructure and 5G networks. The ITU has provided probably the most exhaustive graphic and textual summary of 5G usage scenarios (2018, summarised in diagram below): a) Enhanced mobile broadband (eMBB) – enhanced indoor and outdoor broadband, enterprise collaboration, augmented and virtual reality; b) Massive machine-type communications (mMTC) – IoT, asset tracking, smart agriculture, smart cities, energy monitoring, smart home, remote monitoring; c) Ultra-reliable and low-latency communications (URLLC) – autonomous vehicles, smart grids, remote patient monitoring and telehealth, industrial automation.

There are various estimates about the deployment and economic impacts of 5G which are still not mature and consolidated but provide good illustrations. The new networks are deemed crucial to secure the strategic autonomy of the Union (NIS Cooperation Group, 2019, p. 3). In March 2019 the European Council supported a concerted approach to the security of 5G networks and the European Commission adopted the Recommendation Cybersecurity of 5G network

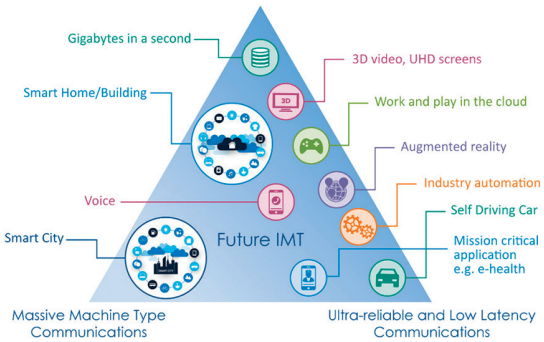


Figure 4: 5G usage scenarios, source: ITU (2018, p. 7)

((EU) 209/534 of 26 March 2019). Earlier the Commission had adopted the 5G Action Plan, setting 2020 as the target for roll-out of commercial 5G networks (European Commission, 2016a). The plan foresaw the deployment of 5G infrastructures and services across the Digital Single Market through a mix of private and public investments.

A more balanced view is, however, also needed for there are various obstacles and challenges (ITU, 2018; ITU 2019). Obstacles include: cybersecurity (see section 2.1.6), spectrum fragmentation, standards development, availability of devices, high CAPEX/OPEX, coverage range (the possible exacerbation of the digital divide already mentioned in the Introduction), and most importantly, the development of use cases that ensure profitable outcomes from the unique competitive advantages of 5G. Particularly noteworthy is the issue of high capital expenditure which generates some scepticism among operators supported by the 5G Infrastructure Association (5GIA), an EU-backed body, and by senior telecom executives cautioning against premature 5G launch announcements (ITU, 2018, p. XII). Last but not least, the deployment of 5G must also consider, and make choices, with regard to the EU's strategic position in the global competitive and geopolitical landscape (Albrycht & Swiatkowska, 2019, pp. 1-3). Critical infrastructures and, especially, military applications resting on 5G/IoT could be disrupted by intentional hostile breaches or may end up being too dependent on suppliers from third countries. A clear example is the government supported expansion of Huawei, to which the US is responding with attempts to prevent this company becoming well positioned in 5G network construction.

2.1.2 IOT

The potential of the Internet of Things (IoT) comes from the combination of hardware architectures (such as sensors, smartphones and wearable devices along with 4G/5G and Bluetooth networks) and software such as "*data storage platforms and analytics programmes that present information to users*" (Walport, 2014, p. 13). The IoT will have a huge impact on automotive, industry, retail and smart building equipment. The large trove of data generated by IoT connections and devices will create fresh resources for growing data analytics and AI in Europe (Palovirta & Grassia, 2019). The opportunities created by ICT are especially evident. IoT will further accelerate networking of individual industries and infrastructure sectors. There are many best-practice examples.

As for the case of 5G, being an emerging and hyped technology there are various very different estimates about the deployment and economic impacts of IoT that are still not mature and consolidated. Some issues, however, require attention or are raising concerns. The first is cybersecurity: the risk of increasing the surface of attack (see more on this in section 2.1.6). Second, there are few business models for profitability (Palovirta & Grassia, 2019). Third, there are not yet any clear 'winners' for interpreting the data from devices or for connecting them to one another. Competing platforms and industrial coalitions are emerging; and there are many different infrastructures that will form IoT networks (Walport, 2014, p. 16). Last but not least, the IoT "is based on the reality that more and more devices will be connected to each other via the internet, allowing data to be shared for analytic processing in the cloud" (Feldstein, 2019, p. 23), which adds to the challenges of adequately protecting personal data.

2.1.3 CLOUDS

Cloud computing can be defined as "*the offer, use and charge for IT services dynamically adapted to demand and supplied through a network*" (BMWi, 2019, p. 5). It encompasses, among other things, infrastructure (e.g. processing capacity, storage space), platforms and software. As cloud computing converges with IoT and 5G, a paradigm shift will take place where increasing volumes of data will be generated and processed (because of real-time needs or because of intellectual property protection and/or data protection) on a decentralised basis.

Interoperability is one of the key advantages of cloud computing allowing implementation of IaaS and SaaS to streamline how IT administrators utilise various hardware and software components from different vendors (Jacobs, 2019, p. 5). For organisations wishing to shift from having data silos to better integrated data gathering and processing, cloud computing would be an essential step towards improving time efficiency and data interpretation. But such a shift may not be possible for all EU Member States, with barriers "*related to data availability, silos, skills, privacy frameworks and impact assessment tools*" (Battisti et al., 2019, p. 10). More so, it is necessary in the first place for administrative agencies to be more open in their sharing of data, and so there needs "*to be clear benefits for public administrations to share their data and there have been pilots that deliver value added for data holders: from quality checks to advanced analytics to GDPR compliance test*"(Battisti et al., 2019, p. 13). It is worth also noting the tension between the technological independence of the location of storage in cloud computing on the one hand, and national regulations on storage locations and data ownership on the other hand (e.g. in health data).

The centralisation of data storage and processing in the cloud at affordable costs means that the benefits from the data economy can be multiplied enabling AI and other applications in healthcare, in the targeted distribution of scarce goods, and through greater resource efficiency. This can generate productivity growth, process optimisation, or innovations in the form of new products and services. Linking up and analysing various data sources opens up additional value creation opportunities, notably thanks to the methods and processes of AI. This development explains why the rapidly-scaling cloud offerings have emerged from the market of large web providers.

A concern for Europe, given implications in terms of data protection and security and of potential economic impact, is that in the cloud market non-EU cloud infrastructure providers currently account for about 80 percent of the global market (Palovirta & Grassia, 2019). This is also underscored in the German document on GAIA-X stressing that "*the existing cloud offerings are dominated by non-European providers with significant market power and rapidly upscaling cloud infrastructures. European alternatives do not offer comparable market capitalisation, scalability or breadth of applications; they are active in specialist niches at best*" (BMWi, 2019, p. 5). Hence, the justification for the launching of Gaia-X as a way to achieve strategic digital autonomy and reduce external dependency.

2.1.4 PLATFORMS

As seen, the growth of platforms has led to worries of data abuse, privacy violation and proper distribution of profit generated by data (Lee et al., 2017). Data governance within platforms, where there are multiple parties contributing, deriving and using data, complicates ownership, access, usage and profit-sharing of collected and derived data. These complexities lead to a larger attack surface and decrease of trust. For a detailed analysis of platforms and the consequences for data flows, see Annex 4.3.

According to Gawer (2009), certain types of platforms can function as the building blocks upon which an array of firms can develop complementary products, technologies or services to innovate. A distinction between intermediation-driven and innovation-driven platforms can be derived from the previous EIT Digital study (2019, pp. 7-8). In that report the following three types were identified:

- **Transaction platforms** facilitate exchange or transactions between different users, buyers, or suppliers. Typical examples are Uber, Airbnb, eBay, and also digital labour markets matching employers and workers (i.e. Upwork, Amazon Mechanical Turk, TaskRabbit).

- **Innovation platforms** facilitate players loosely organised into an innovative ecosystem to develop complementary technologies and products or services.

- **Integrated platforms** facilitate both transactions and the emergence of an innovation ecosystem. The typical example is Apple, which has both matching platforms like the App Store and a large third-party developer ecosystem that supports content creation on the platform. Other examples are Google, Facebook, Amazon, and Alibaba.

We can conclude that some of the players included in the integrated type qualify as platforms intermediating abstract services. On the other hand, none of the dominant platforms qualify as truly open innovation-driven ecosystems. This latter type includes small and emerging ecosystems such as the Ocean Protocol Foundation new platform and SOLID. Large intermediation platforms, however, almost monopolise access to data, as we show in the Annex. This leads us to the implications of a data-driven economy for the long-term economic development of countries.

The US and China dominate the data-rich intermediation layer, whereas France and the UK show up mostly in the production layer. The Top 25 platforms attract most of the visits and, most likely, most of the data. They are thus major economic powers. Lastly, US platforms receive traffic and data from most countries, whereas other countries struggle to keep traffic domestically. Only 22% of the national traffic in France is on national platforms. Overall, most traffic goes to US sites, about a third to national sites, and a tiny portion to sites of third countries (see Figure 21, Figure 22, and Figure 23 in Annex, Section 4.3).

Looking at these numbers, one may ask the questions: Do data flow imbalances make a difference in national economic trajectories? If a country exports more data than it imports (or the opposite), should anyone care? Does it matter what lies inside those exports and imports—for example, 'raw' unprocessed data as compared to sophisticated high-value-added data products? (Weber, 2017, p. 338). In Annex 4.3 we present some examples. During the period 1945-1982, when the Import Substitution Strategy dominated the theory of economic development, the answer would have been that it made a difference, since exporting raw materials and importing finished products was considered as the path to economic decline. In the period 1982-2002 of the Washington Consensus the spread of ICT and reduction in transportation pushed to unbundle supply chains, move pieces behind borders and organise the pieces. Since 2007-2008 the idea that it made a difference went to the background, as global flows of all kinds, except data, have decreased and is not back to pre-crisis level).

In the context of the new data economy there are two arguments for making the case that the question above does not matter. First, that of the absolute gains from data flows claiming that what matters is being part of such flow. The McKinsey Global Institute (MGI) has put forward a very clear articulation of this position, arguing that directionality and content is irrelevant because data flows "*circulate ideas, research, technologies, talent, and best practices around the world.*"(MGI, 2016). The second, supported by Silicon Valley, claims that 'open is best', which has been put forward vociferously whenever the EU has introduced or attempted to introduce regulation of platforms affecting US tech giants (Kennedy, 2015) or introduce the digital services tax (ITIF, 2019, p. 3). A third position, however, is what Weber calls 'data nationalism' as a sort of reflexive response and consists in trying to have their own data value-add companies 'at home' and to stop the new oil to flow abroad for the extraction of surplus (i.e., through data localisation laws or provisions within law). Weber does not embrace data nationalism, yet he argues that a sort of new digital import substitution strategy is possibly

the only alternative for a mid-sized country that is currently in a peripheric position (i.e. exporting raw data and importing data-driven finished products and services). The options for a mid-size country (and may be for the EU as a whole) are the following:

(1) Join the predominant global value chain led by American platforms and seek to maximise leverage and growth prospects within it to catch up.

(2) Join a competing value chain, like Chinese intermediation platform businesses and try to do the same.

(3) Combine (1) and (2).

(4) Insulate or disconnect to a meaningful degree from those value chains, and work to create an independent data value chain within the country or perhaps regionally within the European Union.

The first three strategic options are really variants on one big choice: does joining existing global data value chains point toward an economic and technologically advantageous future? Analysis (see Annex 4.3) suggests a healthy dose of scepticism about that prospect. The catch-up argument would in principle depend on the EU climbing a development ladder that starts with outsourced lower-value-added tasks in the data economy and climbing it at a faster rate than the leading economies climb from their (higher) starting position. Hence, a New Import Substitution Industrialisation in the digital economy may be the only viable position.

Aside from whether or not one agrees with this position, the analysis is a warning on the dangers that European countries become only fields for the extraction of data (the new crude oil) by foreign intermediation platforms. The more data US firms absorb, the faster the improvement in the algorithms that transform raw materials into value-add data products. The better the data products, the higher the penetration of those products into markets around the world. And since data products generate more data than they use, the greater the data imbalance would become over time.

2.1.5 ARTIFICIAL INTELLIGENCE

To delimit and define how we consider AI below we refer to a discussion on the importance of having a definition of AI before introducing regulation. As suggested by Buiten (2019), to avoid circular definitions we limit our analysis to AI underlying technology: machine-learning algorithms. Algorithms are instructions given to computers to follow and implement, in tasks such as ordering possible choices (prioritisation), categorising items (classification), finding links between items (association) and removing irrelevant information (filtering), or a combination of these. Machine Learning (ML) algorithms are more sophisticated as they learn from data. ML is used extensively for a variety of tasks including web search, spam filters, recommender systems, ad placement, credit scoring, fraud detection, stock trading, drug design, and many others (Domingos, 2012). Further advancements have led to the use of deep learning using artificial neural networks that can be applied to a wider typology of data (i.e., voice). The more training data the neural network processes, the more accurately the neural network can begin to process new, unseen inputs and successfully return the right results (Hof, 2013).

AI has allowed for remarkable improvements in numerous fields. But serious concerns have been raised about accountability, fairness, bias, autonomy and due process of AI systems. The roots of biases in ML and deep learning are in data, testing, and decision models used and their analysis debunks some myths. It is not true that big data ensures validity and accuracy; also the quality of data matters (Domingos, 2012). If key data is withheld by design or chance, the algorithm's performance might become very poor (Olhede & Wolfe,

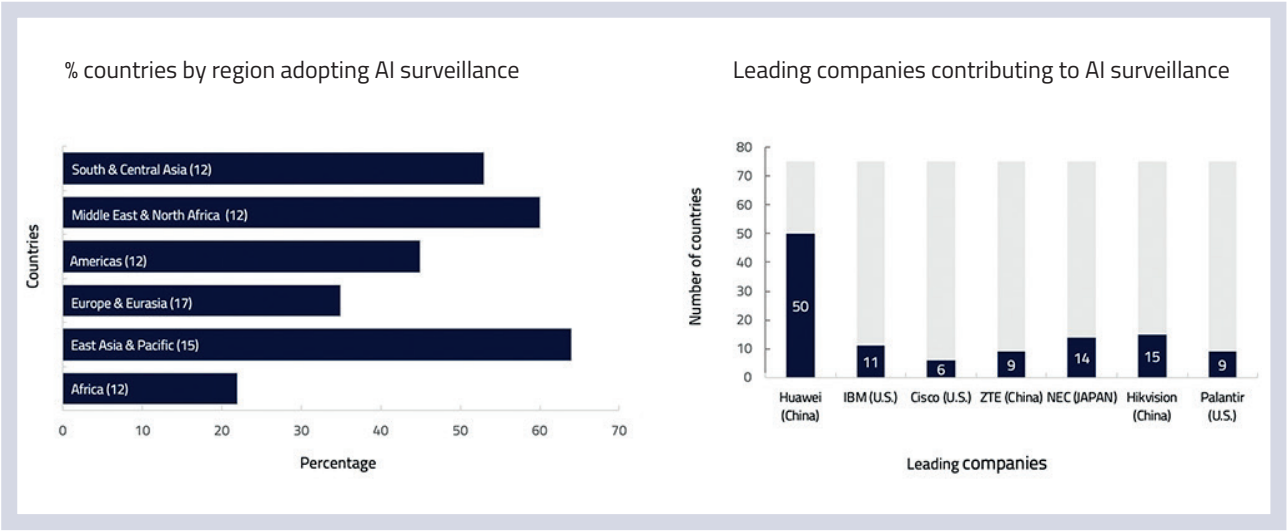


Figure 5: AI surveillance - country adoption and leading suppliers, source: AI Global Surveillance Index (AIGS) reported in Feldstein (2019, p. 9)

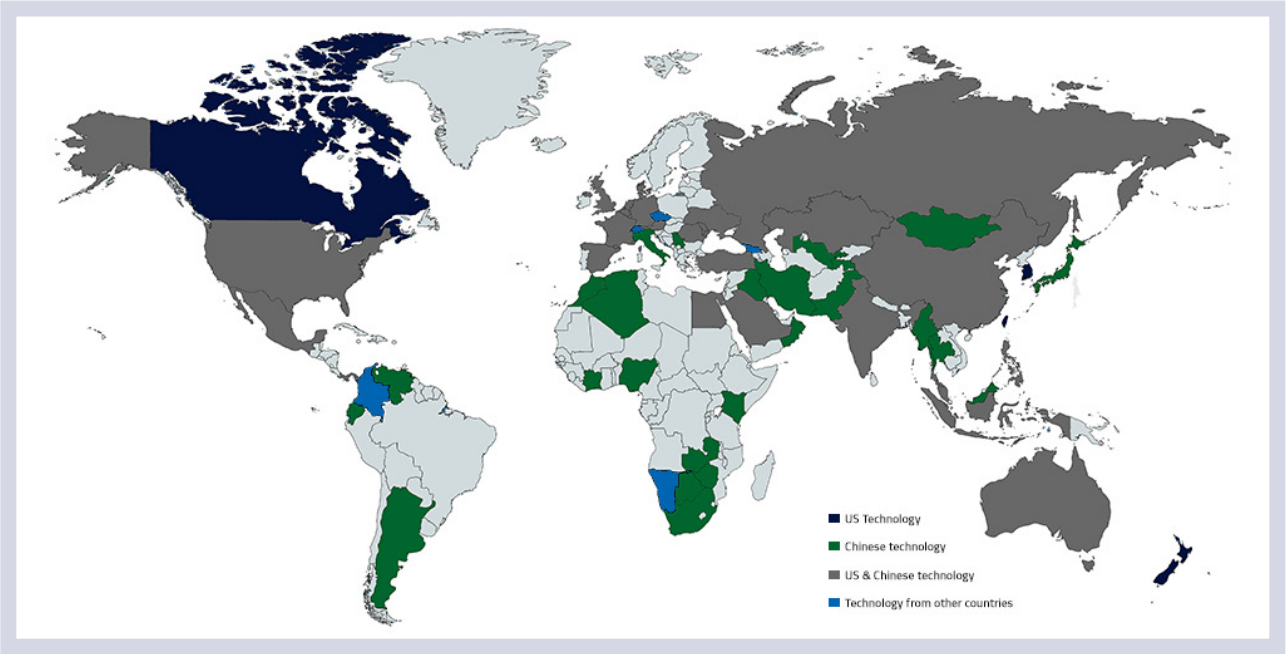


Figure 6: AI surveillance technology origin, source: AI Global Surveillance Index (AIGS) reported in Feldstein (2019, p. 1)

2017). The often-used implicit assumption that once we collect enough data, algorithms will not be biased, is not justified (Barocas & Selbst, 2016). Equally worrying, an increasing number of states are "*deploying advanced AI surveillance tools to monitor, track, and control citizens to accomplish a range of policy objectives*" (Feldstein, 2019, p.1). Chinese companies Huawei, Hikvision, Dahua and ZTE "*supply AI surveillance technology in sixty-three countries, thirty-six of which have signed onto China's Belt and Road Initiative (BRI)*" (ibid.), while US firms such as IBM, Palantir and Cisco supply around thirty-two countries (Feldstein, 2019). And "*liberal democracies in Europe are also racing ahead to install automated border controls, predictive policing, safe cities, and facial recognition systems*" with many "*safe city surveillance case studies posted on Huawei's website relate to municipalities in Germany, Italy, the Netherlands and Spain*"(Feldstein, 2019, p. 8). The distribution of these surveillance technologies is shown in Figures 5 and 6.

A typical take in the media coverage on the issue of regulating AI is that it is either impossible (Spencer, 2019) or not a good choice (Carmeron, 2019). The impossibility argument is the lack of enough knowledge by the regulators and that, ironically, "*Artificial Intelligence regulation may be impossible to achieve without better AI*" (Spencer, 2019). The arguments for not regulating are the typical concerns about stifling innovation, global competition (i.e. China will do it anyway and will overcome us), uneducated regulators, and the mantra "*we have never regulated science*" (Cameron, 2019).

In an article published by the New York Times three general rules on how to regulate AI are presented: a) AI systems must be subject to the full gamut of laws that apply to its human operator. This rule would cover private, corporate and government systems; b) AI systems must clearly disclose that they are not human; c) AI systems system cannot retain or disclose confidential information without explicit approval from the source of that information (Etzioni, 2017).

Both in the US and in Europe some initiatives on AI regulation have occurred. The 2016 White House report on AI suggests that many of the ethical issues related to AI can be addressed through increasing transparency (Executive Office of the President, 2016). Attempts are being made by regulators in the US to improve the transparency and accountability. For instance the Algorithmic Accountability Act that aims to direct the Federal Trade Commission (FTC) with creating detailed policies to ensure oversight for automated decision-making systems (Teich, 2019). Such oversight is meant to take the form of "*Impact assessments where the data and methodology for training algorithms are documented*" (Teich, 2019). The European Parliament in its 2016 report on AI notes that "it should always be possible to supply the rationale behind any decision taken with the aid of AI that can have a substantive impact on one or more persons' lives and "*to reduce the AI system's computations to a form comprehensible by humans*". The main focus of the EU guidelines on AI development are set out in a 2018 Communication (European Commission, 2018) and summarised in a European Parliament briefing document (European Parliament, 2019c). The EU aims to ensure a human-centric approach that is respectful of European values and principles so as to build a trustworthy framework where AI systems are lawful, ethical and robust. A supplement to the European human rights approach is the data perspective advanced by the German Data Ethics Committee (DEK) that focuses on the origin and potential impact that data gathering and processing "*may have on certain parties who are involved with the data, such as by being the data subject, as well as on society at large*" (DEK, 2019, p. 8).

Finally, the General Data Protection Regulation (GDPR) lays down a right for a data subject to receive meaningful information about the logic involved if not only information is collected about them, but also profiling takes place (Arts 12, 14 and 22). In particular Article 22 about the right to an explanation of a decision taken by an algorithm has spurred a debate on the possibility to introduce a legal requirement of algorithm transparency (Buiten, 2019; Goodman & Flaxman, 2017; Wachter et al., 2018). Article 22 may prohibit a large range of algorithms currently in use such as recommendation systems, credit and insurance risk assessments, computational advertising, and social networks. Biases in algorithms producing discrimination, however represent a clear concern, since the right to non-discrimination is deeply embedded in the normative framework that underlies the EU, and can be found in Article 21 of the Charter of Fundamental Rights of the European Union, Article 14 of the European Convention on Human Rights, and in Articles 18-25 of the Treaty on the Functioning of the European Union.

The use of algorithmic profiling for the allocation of resources is, in a certain sense, inherently discriminatory. On the other hand, according to Wachter et al. (2018, pp. 842-843) opening the black box to ensure the right to explanation faces four barriers: a) the GDPR does not contain a legally binding right to explanation; b) it would apply only when a decision is entirely automated and produces legal or other significant effects; c) explaining how algorithms work and go wrong is technically very challenging ; d) data controllers will be reluctant to share details of their algorithms to avoid disclosing trade secrets, violating the rights and freedoms of others (e.g. privacy), and allowing data subjects to game or manipulate the decision-making system. These difficulties, however, do not detract from the importance in terms of social and ethical value of offering explanations to affected data subjects, and the authors propose as an alternative the use of a counterfactual explanation (Wachter et al., 2018, pp. 843-844)[56].

2.1.6 CYBERSECURITY

"*Cybersecurity and digitalisation are two sides of the same coin. This is why cybersecurity is a top priority. For the competitiveness of European companies, we have to have stringent security requirements and a unified European approach. We have to share our knowledge of the dangers. We need a common platform.*" (Von der Leyen, European Commission President).

The rise of cybersecurity threats can be attributed to the expansion of the attack surface determined by nearly universal ICT usage and the take-off of IoT (In Annex 4.4 we provide selective evidence on the increase globally of cybersecurity breaches, their costs, and especially how they hinder the full development of the data economy as well as further details on regulation). In general, we observe an oxymoron with respect to cybersecurity. Everyone knows about rising rates of cybercrime even from the daily news. All companies handle sensitive information in their ICTs which can be a target of a cyberattack. Yet, cybercrime incidents do not seem to decrease, which implies that not enough is being done. The explanation of this situation can be found in the behavioural bias and information asymmetry described in Section 2.1, and to fragmentation in the regulatory approaches both within and beyond the EU.

The major critical issue in the global cybersecurity landscape is the lack of a global cybersecurity viewpoint shared by all nations (much like the United Nations attempts to set common welfare and social standards across the globe) and a global cyberattack resilience strategy[57].

We see specific security issues concerning **5G, IoT, critical infrastructures**, and **AI**.

The main security issues of **5G** are: First, the value chain of 5G includes many stakeholders which is one important source of security risks. Mobile network operators will play a key role, but many other players enter into the picture. Second, related to the previous source, new technical features (a move to software and virtualisation through 'Software Defined Networks (SDN) and Network Functions Virtualisation (NFV) technologies; 'Network slicing'; Mobile Edge Computing) bring new security challenges, increasing the complexity of the supply chain. Third, functions currently performed physically and logically separated will move closer to the edge of the network. If not managed properly, these new features are expected to increase the overall attack surface and the number of potential entry points for attackers.

With **IoT** there is danger implicit in increasing interconnectedness of hardware and software utilised by businesses and governments in their adoption of IoT technologies and applications (Walport, 2014, p. 20). As the value of the IoT is the data extractable from its functions, any security weaknesses that are exploited can be economically and personally costly (i.e. in terms of loss of important assets or infringement of privacy). The new risks of IoT apply in particular to consumer IoT, as it can involve 'non-technical' or 'uninterested' consumers, who connect an increasingly wide variety of devices to their home networks[58]. It also relates to the fact that various sectors and industries heavily depend on ICT components and on interdependence between current and future infrastructures (e.g. in smart cities environments, connected cars, energy smart grids).

As summarised by the European Parliament (European Parliament, 2019a), there is a specific need to focus security efforts on **critical infrastructure**. Energy and other utilities are increasingly controlled and monitored by networked industrial control systems. The electricity grids are being transformed into smart grids, in which more and more control functions are automated. This is expected to grow with the full deployment of 5G and IoT, which as seen also potentially in-

crease the risks[59]. The European Commission has developed various activities on critical infrastructures since 2006 (see Annex 4.2).

Finally, there are also cybersecurity implications for AI as it plays a role in risk management of cybersecurity (Timmers, 2019a). What are the ethical challenges in cybersecurity risk management, notably when making use of AI? Extensive monitoring and pervasive risk-prevention with the help of AI can be highly intrusive and coercive for people, whether employees or citizens. AI can also be so powerful that people feel that their sense of being in control is taken away. They may get a false sense of security too. Deep-learning AI is, as of today, not transparent in how it reaches a decision from so many data points, yet an operator may blindly trust that decision. AI can also incite freeriding, as it is tempting to offload responsibility onto 'the system'. We are therefore confronted with a plethora of ethical issues when combining AI and cybersecurity in a risk management approach to strategic autonomy. They include erosion of individual autonomy, unfair allocation of liability, the fallacy of human-in-the-loop, the contestable ethics of mass surveillance and of trading off individual casualties versus collective protection.

Cybersecurity markets vary from monopolistic to competitive and fragmented structures. In China, the cybersecurity market is dominated by large monopolies with links to the national security apparatus (Cheung 2018). In Japan the role of the Ministry of Economy, Trade, and Industry (METI), as well as a long-established practice of top-down policymaking have contributed to the slow speed of growth in the cybersecurity sector (Bartlett, 2018). In the United States there is a plethora of companies marketing their cybersecurity programmes (Aggarwal and Reddie 2018b). In Europe markets are fragmented between a few large players and several small firms (Carr & Tanczer, 2018; D'Elia, 2018; Griffith, 2018; Timmers, 2018). Such structures are to a large extent shaped by the fact that national governments intervene on the basis of national security concerns. This is documented for the US (Aggarwal and Reddie 2018b), China (Cheung 2018), Finland (Griffith, 2018), France (D'Elia, 2018), and for EU as a whole (Timmers, 2018). In this respect, Timmers notes that, in many European countries, cybersecurity suppliers developed through a close relationship to military and government buyers. The downside is a degree of national institutional dependency: "*Historically, industrial development in this area has been stimulated by governmental procurement and some highly innovative European companies in this sector are still largely dependent on this in their home country. A side effect of this situation is limited willingness for cross-border procurement, which is a barrier to the development of a common cybersecurity market*"(European Commission, 2016b). Both the United States and Europe are experiencing shortage of programmers and computer scientists working on cybersecurity issues (Aggarwal & Reddie, 2018b; Carr & Tanczer, 2018; Timmers, 2018), and in Europe there is also a lack of capital to fund innovation and market growth (Timmers, 2018).

The EU cybersecurity policy has been developed in response to three drivers: preserving the internal market, combating terrorism, and playing a global role (Timmers, 2018). It started in 2013 when a fully-fledged EU Cybersecurity Strategy was launched, and a landmark EU cybersecurity law focused on economic resilience was proposed: Network and Information Security Directive (NIS Directive)[60]. In 2016 an EU private-public partnership on cyber security increased investment in research and innovation. Driven by the rapid rise of cyber incidents, this can be characterised as a step towards a comprehensive and integrated EU cybersecurity policy. Currently, we are in a phase which started in September 2017 with an ambitious renewal of the overall strategy and several important legislative proposals. These include the EU Cybersecurity Act which introduces EU-wide IT security certification and an extended mandate for the cybersecurity agency ENISA, legislation for a common approach to scrutiny of foreign direct investment including for cybersecurity concerns, and legislation for strengthening EU cybersecurity competence. An EU meeting of all Heads of State also discussed cybersecurity.

This third phase can be characterised by cybersecurity becoming a top political priority. The most recent EU industrial policy argues that industry should become more adaptable, innovative and open to digitisation in order to be globally competitive. EU cybersecurity industrial policy is thus firmly embedded in general EU industrial policy. The success or failures of industrial policies and of other types of policies depend on many factors in the design or implementation, including the possibly unforeseen strategic behaviours of the actors subjected to the policy. An interesting case is that of breach notification laws aimed at market modification in the direction of incentivising firms to invest more in cybersecurity in order to avoid having to publicly report about the breaches. A quasi-experimental empirical study of the effects of California's law (introduced in 2002) found that while data breach notification laws have received considerable attention in recent years, their impact on firms' investment in web server security appears modest (Murciano-Goroff, 2018)[61]. Yet, the Californian law did not include heavy fines. A theoretical principal/agent model shows that breach notification laws can produce social benefit (enough cybersecurity investments by firms to have positive overspill on economy and society) only if the fines foreseen are large enough (Laube & Bohme, 2016)[62]. This would suggest that the European GDPR and NIS Directive, both of which include sizeable fines, may be more effective than laws adopted by states in the US.

2.2 Data protection

2.2.1 INTRODUCTORY OVERVIEW AND KEY CONCEPTS AND DIMENSIONS OF ANALYSIS

As of today, 107 countries (of which 66 are developing or transition economies) have established regulations for protecting people's data and their privacy. Notably, Asia and Africa demonstrate similar data regulation adoption, with less than 40 per cent of their countries having a relevant legislation[63].

Globally, there is an increasing growth in data protection laws, many of which have been modelled on comprehensive guidelines or regulation such as the EU GDPR, or the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. As mentioned earlier, according to UNCTAD Data Protection Tracker 4, over 100 countries around the world now have data protection laws in place. In Figure 7 is a summary of which countries across the globe have full or draft data protection legislation in place, based on this tracker. The regulatory landscape on data protection is presented in Annex (Section 4.2); below we discuss the area of data governance which is directly affected by regulation on data protection. We adopt a more analytical perspective with, however, several references to the current situation and also to possible alternatives. We come back to regulation with some general considerations at the end of this section.

2.2.2 DATA GOVERNANCE AND DECENTRALISATION

Governance involves the allocation of authority to certain parties empowering them to make decisions and influence behaviour[64]. When it comes to data governance, the power and the decision-making involves data resources. The main objective of data governance is to put in place roles and processes to ensure that the data assets of an organisation can be effectively used by the organisation to fulfil its mission; this involves policies and standards, as well as monitoring mechanisms. Consequently, data governance policies are informed by regulatory frameworks such as the GDPR. The power to shape, apply and safeguard policies rests with organisations, individuals and the people or systems that act as their agents. Data governance policies can dictate who has access to what data, how such data can be used by which party, whether data can cross systems or borders, how data quality is established, what processes are in place to ensure data integrity, what happens when there is a data access breach or when data quality is compromised.

Bishop (2017) points out that data governance in Europe faces a number of challenges. Concerning data protection, the following issues are highlighted. Firstly, the definitions of private and privacy are ambiguous (Bishop, 2017, p. 4). This ambiguity is also present in defining whether social media platforms are public or private, especially as users may believe that social media platforms are private from reading user agreements, but this is not always the case. More so, the data costs and analytical complexity driving collaboration between makers in public and private organisations is blurring the distinction between public and private use of data. This blurring of distinction between public and private use, affects how data science classifies their research – if they infringe on the privacy of data subjects then this is a human rights matter, but if privacy is not infringed then their research may be exempted and not be classified as human subjects research.

In the global data ecosystem, a related question that also arises is who has the authority to make, monitor and enforce data governance policy, especially when it comes to data protection. In *centralised* systems there is a single party (organisation or individual) on whom this authority rests, while in *decentralised* systems authority can rest on different parties. Decentralisation introduces benefits in terms of scalability but also requires mechanisms of establishing trust and harmonisation between the decisions that different parties make; computer intermediation can help address some of those issues. One can argue that fully decentralised systems, on all grounds, are often as impossible as fully centralised ones. In most cases, decentralisation proposals have focused, as a first step, on: a) data container ownership and access control; and b) identity management. The next few paragraphs deal with these two dimensions of analysis, without considering the GDPR, the latter is discussed together with other aspects in Section 2.2.6.

2.2.3 DATA OWNERSHIP AND ACCESS CONTROL

In terms of container ownership and access control, in a centralised approach data can be stored in a database owned and/or controlled by a single party; even if data can be distributed over clusters of computers, the data stores are owned and controlled by that party. In the case of data ecosystems such as that of Facebook, access to some of the data can be provided to third parties (application developers). Users have some moderate control over their own data on such platforms depending on company policy and regulation. From a data governance viewpoint, even partial rendering of control to the users or to third parties (with users' consent) is an act of decentralisation. Nevertheless, in mainly centralised data ecosystems it is not clear (i) whether the maximum potential of innovation has been exploited; and (ii) how equitable is access of third parties to such data ecosystems (with users' consent). These questions are particularly important for Europe given that most of those ecosystems are owned and controlled by companies in the US or China[65].

On the other hand, an example of decentralised data go-

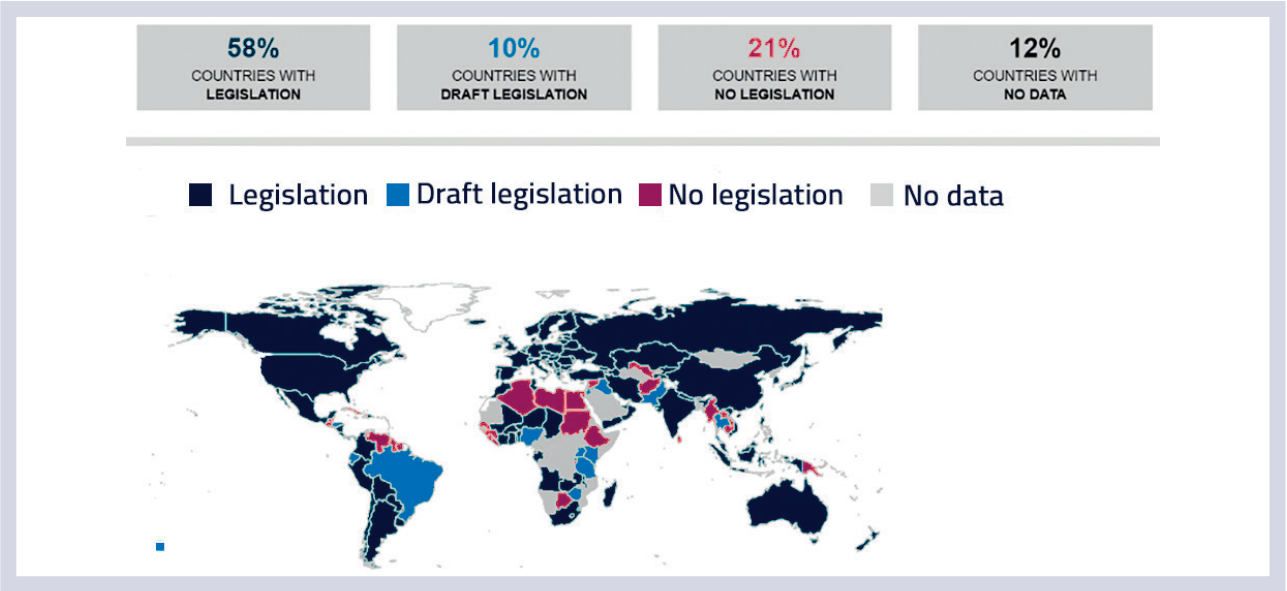


Figure 7: Data Protection and Privacy Legislation Worldwide, source: UNCTAD (2019)

vernance in terms of access control is that proposed and implemented in the SOLID[66] framework, created by Tim Berners-Lee, the inventor of the World Wide Web. The first principle of SOLID is the separation between application logic and data; i.e. application logic and data stores (containers) can be owned and controlled by different parties. In SOLID environments a user can own and control their own personal data store (or POD for short), while application developers can add value to those data with the consent of the personal data store owners (the users). Sections 2.1.3 and 2.1.5 address the potential of AI over data stored collectively on the cloud, which presents a possible tension between decentralised data storage and the innovation that AI can bring on centrally-stored data. Nevertheless, PODs can be collectively stored on POD Servers on cloud infrastructures, which can partly address this tension if a large number of users authorise a third party with access to their cloud-stored PODs.

One other issue in this form of decentralised data governance is the need for a framework to establish trust between parties so that they can agree to provide each other with access to their data containers. It has been argued that technology solutions like Blockchain could address some of those issues to some extent[67] but there are concerns (and proposed solutions) to the scalability (James-Lubin, 2015) and to the energy requirements of some forms of blockchain (Orcutt, 2017).

2.2.4 DATA OWNERSHIP AND ACCESS CONTROL

Another fundamental aspect of data governance is that of identity management for individuals, organisations and resources involved. Identity is key to putting access control policies in place, for identifying data assets and where they are stored, and for establishing trust between parties. In a centralised identity management system, a single party has the authority to serve as the root of trust for identifiers, verifying that a party in possession of that identifier is who they claim to be[68]. Trust management systems can also control the allocation and management of identifiers. In enterprise systems, Identity Management or Identity Access Management (IAM) is part of the data management frameworks and the authority to assign identifiers rests with the enterprise. Enterprise systems by companies such as IBM include IAM components. Cloud platforms, like Amazon Web Services[69] and Microsoft Azure[70], provide support for IAM functionality to their customer enterprises, which is often called ID-as-a-Service (IDaaS). Across enterprises, on the macro scale of the Internet, identifiers involve domain names that are assigned and verified using a hierarchical structure, the top level of which is centrally managed by ICANN, while the lower levels are managed by different organisations, businesses and individuals. When it comes to online interactions on the Internet, concerns over centralised identity management systems include:

1. Individuals rely on a central authority to assign and verify their identity online.

2. Individuals rely on a central authority to keep safe their personal information that is used to verify their identity.

3. A central authority can monitor the online transactions of an individual based on the identity verification requests that it receives for that individual.

Such concerns have led to proposals for self-sovereign identity (SSI) and standardisation by the Decentralised Identity Foundation (see Annex 4.2.4).

2.2.5 DATA PROCESSING

Alongside a more stringent policy framework, taking stock of data architectures and their economic value is also necessary. A collection of relevant data storing, and processing technologies, platforms and marketplaces would be very large. 'Big data' technologies established the feasibility of storing and processing data on a large scale and concerned primarily systems with centralised data governance. The various data-collecting and processing platforms, as described in Section 2.1.5, were developed with the aim to process large, centrally collected and controlled datasets. However, the data economy is also showing potential to develop towards decentralised data ecosystems, given increased calls for data protection regulation and the affordances of emergent frameworks outlined above. Frameworks like SOLID and platforms such as HATDEX[71] provide solutions for application deployment on decentralised data stores. Such frameworks show a trend towards systems that can develop and leverage the 'long-tail' of data assets, which are not collected and managed by big corporations. This could present a significant opportunity for European businesses.

2.2.6 FINAL CONSIDERATION ON GDPR AND BEYOND

In the domain of data protection, the GDPR is being very influential and continues to be highly debated, but it remains a first step[72]; advocates of decentralised models and subject sovereignty call for more as in the Declaration of digital independence by co-founder of Wikipedia Larry Sander[73]; others criticise it as either unfeasible or as potentially stifling innovation and further exacerbating the problem of fragmented data markets in Europe because of its rigid regulation of data sharing (The Economist, 2018b). From both legal and technical perspectives, the right to withdraw consent, the right to be forgotten, and the right to explanation remain controversial both in terms of their feasibility and their potential disruption of existing practices and business models (Buiten, 2019; Goodman & Flaxman, 2017; Li et al., 2019; Politou et al., 2018; Wachter et al., 2018). For instance, the right of be forgotten – if implemented – affects the value of already-stored data and would impose quite a burden on controllers who need to inform a vast number of third parties "*when a data subject has requested the erasure of previously published personal data relating to them*" (Politou et al., 2018, p. 12). Articles 13 and 22 requiring that in some cases algorithmic decisions are reviewed and explained may increase labour costs, thus affecting the development of AI possibly increasing its costs and reducing the application scope (Li et al., 2019, pp. 3-4). As seen earlier, the applicability of Article 22 is technically daunting (Buiten, 2019; Goodman & Flaxman, 2017; Wachter et al., 2018), and counterfactual explanation has been proposed as a more pragmatic alternative, performing the same function without opening the black box (Wachter et al., 2018).

We highlighted some of the potential problems with the GDPR. But some of these may require more action. Countering the typical argument that regulation is the mortal enemy of innovation, it is clear that with no regulation the fact that data stored on central repositories are owned and controlled by a few companies limits the innovation potential of data, since fewer AI innovators have access to the trove of data needed to train their algorithms. Emerging evidence on the effect of decentralised systems points to two aspects. First, decentralisation can favour reaping the benefits of the long tail of more specific and sectorial data not yet monopolised by tech giants. Second, these experiences are leading to increasing demand that individuals keep in control of access to their data. This suggests that, if available, individuals may start storing their data securely in personal data store platforms or infrastructures; this may also enable them to receive services on the data that they themselves download: following regulatory and corporate policy developments, Europeans can download their data from big platforms but there is no support for further value to be added to it. It is possible, in accordance with existing regulation and the GDPR, to promote the development of technologies/platforms that enable users to share their data with entrepreneurs who can provide innovative value-added services.

2.3 Critical issues

A number of critical aspects anticipated in the Introduction have emerged more clearly in the previous sections. Some are specific to the digital infrastructures reviewed in Section 2.2, whereas others are more general and actually show how our two dimensions of analysis (digital infrastructures and data protection) overlap and are not fully orthogonal. We start from the specific issues (following the order used in Section 2.2) to move to the more general ones.

The challenges specific to 5G networks have been discussed in 2.1.1 and 2.1.5. They include: potential high capital investments needed; lack of clear use cases; models of profitability; the risk of new forms of the digital divide; and more technical obstacles related to spectrum fragmentation, standards development, coverage range and availability of devices. Most importantly, however, 5G networks raise potentially new security and eventually (after full deployment) data protection challenges; this may be aggravated when 5G is used in combination with IoT. In the case of the latter, the critical aspects are similar to 5G for what concerns security and data protections issues and, while capital investment is less of an obstacle, lack of standards and of business models is. Cloud computing also presents challenges when fully integrated seamlessly with 5G and IoT (and considering use of AI algorithms). With full deployment of 5G and IoT a paradigm shift can be envisioned with Edge Cloud computing (BMWi, 2019, pp. 5-6). This is a decentralised data architecture principle enabling data processing not only in the cloud but also where it is generated. The IoT will generate an increasing and poten-

tially huge amount of data on a decentralised fashion (with sensors or wearable devices). In some cases, real-time processing where a few milliseconds of reaction time (latencies) will be possible with 5G and Edge Cloud computing. Decentralised processing may also be needed for a matter of complying with intellectual property and/or data protection. So, future cloud developments could contribute in combination with 5G and IoT to exacerbate security and data protection challenges, given the huge amount of new data that could be processed both in centralised and decentralised fashion.

For AI algorithms one of the key challenges is how to address their potential biases and their discriminatory effects. Issues of accountability and auditability of AI software in producing biased or inaccurate results have been raised (Knight, 2017); the quality of data feeding those algorithms has consequently been questioned, not so much on 'data are right' grounds but on grounds of bias and scope (Redman, 2017). Also important is the widespread adoption of AI surveillance tools and the role AI plays in cybersecurity. There are other broader issues for AI algorithms, but these are part of the discussion of more general challenges.

Aside from very specific issues, the many platform and network infrastructures, especially in a scenario of full integration and convergence, point to broader data protection and security challenges as well as implications for Europe's strategic digital autonomy.

The data economy thrives on personal or sensitive data, and this is a major source of concern. In addition, there are those data Zuboff (2019) calls 'behavioural surplus' that users do not even realise are collected about them. In particular, "loss of control over personal information creates a variety of near-term and longer-term risks that are difficult for individuals to understand - and, importantly for antitrust purposes, therefore impossible for them to value" (Cohen, 2019, p. 175). Though legally behavioural data may not be defined as sensitive or personal (but article 3 of the GDPR goes into the direction of expanding the definition of what is personal data), it enables companies to 'hyper-nudge' consumers. Often the methods of obtaining user consent have been arcane and the ethics of processing and use of personal data have been questioned on various grounds (Bishop, 2017). The online marketing practices based on big data analytics have been defined by Yeung as 'hyper-nudges' that guide decision and in practice reduce the autonomous decisions of consumers (Yeung, 2017).

How industry is harnessing big data to transform personal digital data into economic value, has been described by one leading cyberlawyer as the latest form of 'bioprospecting' (Cohen, 2012, 2015). Concerns over feedback loops based on surveillance of online users have also emerged (Zuboff, 2019). The Silicon Valley rhetoric of the Open Internet has it that privacy may be a good that most people are willing to trade away and that a tort-based approach would suffice. Yet, behavioural scholars have amply documented that the 'notice and consent' is a fiction since individuals face insuperable challenges to truly give an informed consent (Acquisti et al., 2015); most people neither read nor understand online privacy policy and, if they read all those encountered, they would spend 244 hours per year or 76 days per year at an opportunity cost that is worth billions (McDonald & Cranor, 2008). Struggling to manage their privacy relations with the hundreds of digital service providers that they interact with online, users find it difficult, if not impossible, to assess the risk of harm in a series of isolated transactions given that many privacy harms are cumulative in nature (Solove, 2013, pp. 1890-1891). It is a domain riddled by complexity, resulting in information asymmetry exacerbated by bounded rationality and behavioural choices. As shown in Section 2.1, users as individuals are aware and concerned, but still are not always capable of making the best decisions. One conclusion from the above may be the acknowledgement that, contrary to the current narrative of privacy as a personal good, it must be accepted that privacy is also an important public good. Citizens should not be free (or be protected) to sell their data as this data can damage others (very similar to pollution).

Cyber breaches are a double source of concern, firstly for the damage they cause and secondly because they generate geopolitical tensions. The damages and the geopolitical tensions hinder the full development of the data economy and may stifle global innovations and knowledge exchanges. The US has been restricting Chinese Foreign Direct Investment (FDI) in several strategic technologies; the EU adopted a measure to monitor FDI (European Commission, 2017b) and set to revise its cyber security strategy to "build greater resilience and strategic autonomy" (European Parliament and Council, 2017, p. 2); China has already taken several actions and made several statements to mark its digital sovereignty (Cheung, 2018). From spear phishing and distributed denial-of-service attacks (DDoS), to advanced persistent threats (APT), cyberattacks have become increasingly commonplace as connected technologies have become ubiquitous (Aggarwal & Reddie, 2018a, p. 291), and some nations have taken a more activist approach than others (Abelson et al., 2015). This may further increase the national and international fragmentation preventing common grounds to emerge to favour innovation and market expansion. Cybersecurity is also characterised by information asymmetry, bounded rationality and behavioural biases on the side of firms, which are compounded by opportunistic behaviours (liability dumping and free riding).

The oligopolistic access to valuable user data by few companies is a concern and an actual barrier to innovation and economic growth more generally. Analysis of traffic data shows that a few US and Chinese tech giants are dominant in having access to data (Faravelon et al., 2016). This relates to two issues: security and innovation. Non-state actors, including platforms, represent a cyber challenge to the traditional state-based system of international relations (Timmers,

2019a). Platforms have cyber power as they hold sensitive information and unmatched potential for surveillance. A report for the European Parliament shows that platforms have become both opponents and partners of governments which try to enforce security through means such as surveillance (Garcia et al., 2014). The innovation barrier is well explained in a technical white paper by the Ocean Protocol Foundation (OPF) where it is affirmed that the greatest beneficiary of the ongoing data driven transformation: "*... are companies that have both vast data and internal AI expertise, like Google and Facebook. In contrast, AI start-ups have amazing algorithms but are starving for data; and typical enterprises are drowning in data but have less AI expertise. The power of both data and AI — and therefore society — is in the hands of few*" (OPF, 2019, p. 5). This means that the potential embedded into new data analytics is only partially exploited due to concentration in access to data.

The above observation is in line with the fact that we are seemingly experiencing another 'productivity paradox' or 'riddle'[74]: despite the widespread hype about AI, its contributions to productivity seem to have been limited thus far (Gordon, 2016, 2018; Nadella, 2017). A potential explanation relates to the role of data. The free data model has made productivity-related data much less accessible than consumption-oriented data. Workers who expect to be compensated are the primary performers of productivity-related tasks and these often occur within firms unwilling to surrender their proprietary internal data to AI companies for free. So, the monetary value of user-contributed data (The Economist, 2018a) can change the balance of how users perceive the benefits of engaging with online platforms; this article takes inspiration from the work of labour economists proposing to treat data as labour by paying users as 'data labourers' (Arrieta-Ibarra et al., 2018).

Lastly, European countries are dependent on dominant foreign platforms in the emerging intermediation economy (Faravelon et al., 2016). As extensively discussed in 2.1.4, traffic and data mostly go from Europe to US platforms that benefit from the extraction of values and behavioural surplus (expression used in Zuboff, 2019) from such data. If strategic autonomy concerns capacity to shape one's longer-term future in the economy, society and their institutions, then a look at the logic of the data economy and dependencies is inescapable. Being an exporter of raw data (for free) and an importer of services (for a price) may not be irrelevant for future economic development (S. Weber, 2017).

3. FROM SCENARIOS TO SMART POLICY

After the overview and discussion presented in Chapter Two, it is now clear how the complexity of the topic at stake is a big challenge for any conceptual simplification and explains the current confusion that characterises debate in the media and policy circles. Digital infrastructures and personal data protection are not orthogonal but rather present clear overlap as they are inextricably and closely related. The same applies if we consider security and personal data protection, which makes cybersecurity an underlying and horizontal dimension of analysis. Data issues are also difficult to disentangle from matters of competition in the data economy and can be linked to labour market issues (i.e. as in the proposal for treating 'data as labour'). All these domains converge with industrial policy and RTD policy where public regulation, public investments and incentives, and a leading public role in standardisation may be important. Last but not least, most of the discussed topics are currently entangled in international tensions and get coloured with considerations on technological sovereignty and strategic digital autonomy.

Nonetheless, conceptual simplifications are still needed to make a complex reality more intelligible and digestible for reflection on policy. Hence, in the overview of possible policy responses below we stick to the two dimensions proposed in Chapter One, but with a slight modification. In section 3.1.2 under digital infrastructures we consider 5G, IoT, cloud computing, and platforms, whereas in section 3.1.3 we discuss AI under policy responses addressing personal data protection; we treat cybersecurity separately (in section 3.1.4) given that it is a horizontal underlying dimension of our analysis.. These paragraphs are preceded by a more general discussion of policy discourses and approaches (in section 3.1.1).

3.1 Possible policy responses

3.1.1 GENERAL APPROACHES: BETWEEN 'LEAVE IT TO THE MARKET' AND 'MAKE IT A UTILITY'

Among the various sources reviewed for this and the previous study dealing with the platform economy (EIT Digital, 2019), one key discourse in the free-market leaning media and think tanks is that any attempt to regulate the current digital transformation would stifle innovation and produce undesirable side effects. In extreme fashion this discourse can be summarised with the view that regulation is the mortal enemy of innovation (Cohen, 2019, p. 178). Hence, for the sake of economic growth and innovation, matters should be deregulated and/or their governance should be devolved to the private sector through various forms of self-regulation and de facto standardisation. A corollary of this discourse is that attempts at regulation are touted as new forms of protectionism. A second discourse, seen especially with regard to AI, takes the form of an 'impossibility statement'. Regulation of current development is and will remain technically complex and beyond the reach of the cognitive tools and processes available to regulators.

The first discourse can be countered on the basis of both historical and economic reasoning. Historically, it has been amply demonstrated that markets are never able to create by themselves the legal and institutional basis needed for their functioning and that the Great Transformation from rural to industrial society was to a large extent made possible by the institutional innovation produced by the state (Polanyi, 1957). Similarly, the state has historically provided the basic infrastructure for economic development. In all situations where "*private industry could not or would not act, the public sector would provide the physical roads, ramps, and rails over which the traffic of commerce could move*"(Deloitte, 2017, p. 7); and this applies also to some of the pillars on which the current digital transformation rests and which are taken for granted (Mazzucato, 2015). Many of the components included inside smart phones and the GPS technology exist thanks to very large public sector investments. Using a free-market approach to the current digital transformation in practice is not neutral, for it is not the same as letting markets self-regulate through the dynamic interaction of demand and supply. First, this would maintain intact those uncertainties that delay innovators and fuel the regulator-innovator dilemma described in the Introduction. Second, it is not neutral in that it would de facto reinforce and crystallise current trends and situations of market power that distort competition and impede new and more distributed forms of innovation (see report of the UK Digital Competition Expert Panel on 'Unlocking Digital Competition' and of the German Data Ethics Commission). Third, protectionism can be the result of both over-regulation and under-regulation. While some of the recent European regulatory initiatives and plans are introduced also with the transparent aim of increasing the competitiveness of European industries, it is also evident that US positions on data protection and anti-trust "*have permitted a race to the bottom in the accumulation of platform power and that the relative US laxity has disadvantaged*

European Internet businesses" (Cohen, 2016, p. 382; Cohen, 2019, p. 178). On these grounds, the view juxtaposed to the approach 'leave it to the market' is that the current digital transformation requires a bottom up rethinking of competition and public utility regulatory regimes (Cohen, 2019, p. 200). We briefly review the debate on public utility, after discussing an alternative view to the second discourse about the 'impossibility statement'.

The impossibility statement implication is that in the age of algorithmic governance emerging as a new form of business strategy, regulators cannot keep up and should only hope and wait until algorithms improve and better self-regulate themselves. The most sustained counter argument has been developed by law scholar Juliet Cohen who makes an analogy between the case of Volkswagen's defeat device and the regulation of the data economy (Cohen, 2016; Cohen, 2019, chap. 6). The Volkswagen case shows that the important issue of reducing automotive emissions forced regulators to enter the domain of "algorithmically driven control and marketing by industry" (Cohen, 2019, p. 171). It is worth noting also that Volkswagen justified the installation of the defeat device to improve engine performance and maintain its reputation as an innovator. This reveals a striking resemblance with the arguments of Google and Facebook for not opening their black boxes and not to be constrained in their steady process of experimentation and innovation. In the same way as for the regulation of automotive emissions, Cohen argues that the current digital transformation requires regulatory innovation not only on the 'what' (new rubrics of activities needing regulation) but also on the 'how', meaning entering the domain of algorithmic governance (2019, 812-185 and 200-201). This requires regulatory innovation in the form of the creation of new institutional mechanisms and technical capacities for defining obligations and overseeing compliance. Furthermore, given the complexity of this undertaking, there is a need to move from a risk perspective backing a cost-benefits approach, to policy and regulation from an uncertainty perspective backing a precautionary approach.

The possibility of imposing common carriage/public utility requirements in the digital ecosystem was first proposed in the US in the context of the debate on net neutrality. This debate about net neutrality (i.e., the obligation to treat all content, sites, and platforms equally), was fuelled by Federal Communications Commission's (FCC) landmark and controversial 'network neutrality' 2015 Open Internet Order[76] with opposing views among law scholars (Candeub, 2018; Cohen, 2016; Yoo, 2018). After the FCC was invalidated jurisdictionally, advocates have embraced common carriage and public utility role as the legal basis to defend net neutrality (Yoo, 2018). According to Cohen (2016, p. 379), the net neutrality debate is the occasion to consider more broadly how to adapt the industrial era notions of common carriage and/or public utility provision to the networked information age and the extent to which regulation of the digital world should incorporate public access and social justice considerations.

As in many other digital domains, also on net neutrality the two warring discourses described earlier face each other. On the one hand, industry players (in this case telecom operators) ask the freedom to experiment with new premium service business models for the sake of innovation, which any regulation would stifle in their view. On the other hand, consumer advocates and small Internet companies respond that price discrimination in the context of closed and dominant platforms threaten distributed and decentralised innovation and freedom of expression. An additional critique to the business perspective is that price discrimination of traffic and users would provide dominant players even more data to extract what Zuboff calls 'behavioural surplus' (2019). The debate on net neutrality boils down to whether "*regulatory institutions should be designed to promote enhanced public accountability or whether instead they should take on configurations more responsive to informational capitalism's needs and goal*" (Cohen, 2016, p. 380).

A technical critique to net neutrality is that the Internet does not show the characteristics that have historically justified a common carriage regime, namely: (1) commodity products, (2) simple interfaces, (3) stability and uniformity in the transmission technology, (4) full deployment of the transmission network, and (5) stable demand and market shares (Yoo, 2018, p. 991). Recently, the issue of imposing a common carriage/public utility regulatory regime resurfaced in relation to dominant online platforms. As reported by the Economist, Elizabeth Warren, a leading Democratic contender for America's presidency, has proposed: a) to unwind anti-competitive tech mergers such as Facebook's acquisition of WhatsApp; and, especially b) that online marketplaces which generate annual global revenues of more than $25bn be declared 'platform utilities' and prohibited from both owning a platform and doing business on it (The Economist, 2019). The rationale for the latter more radical proposal is that "*tech titans are mostly two-headed beasts. They not only operate a market but compete in it too. Amazon owns the world's biggest e-commerce marketplace and also sells products on it under its private labels*" (Ibid.). Hence, conflict of interest may distort the ranking of products and services returned in search hits. Earlier in 2018 Representative Steve King (R-Iowa) dropped the bombshell proposal of converting Facebook and Google into public utilities (Constine, 2018). Others suggest that Facebook, with its increasing influence on the political process, is de facto becoming a public utility (Susarla, 2018). Social media as public utility is already an entry in Wikipedia for almost a decade and is constantly updated[77]. Making social media websites as utilities would require government regulation of various platforms, and the argument for it is that by now they are essential social services and need a more equitable regulation. Public utility regulation for social media has been criticised because it would produce undesirable and indirect effects. The opponents argue that social platforms are not essential as water and electricity, platforms change every year and, last but not least, imposing public utility status may have the counterintuitive effect to lock in a real

monopoly, ending the innovations that large online platform produce with effect also on prices. These neo-liberalist arguments are put forward, for instance, in a report by the US free market leaning think tank Mercatus (Thierer, 2012).

3.1.2 DIGITAL INFRASTRUCTURES

If we consider the possible future integrated development of 5G, IoT, and Edge Cloud computing, a number of considerations made in Sections 2.1 and 2.3 can be brought together here that may justify a stronger role of public authorities and consideration of treating this combination as a new public utility digital infrastructure.

As explained in Section 2.1.1, the deployment of 5G networks may be hindered by very high capital investments and uncertainty on profitability. This may discourage operators to deploy 5G at all or may lead them to do so only in profitable densely populated urban settings with the risk of excluding sub-urban and rural areas. IoT deployment may be hindered by lack of standards and of clear business models and uses cases (Section 2.1.2). With many different types of communications mechanisms and protocols, it is yet unclear which will be the foundational infrastructure of IoT. Additionally, with the future growth of connected IoT devices in future years there will be a strong pressure on the allocation of existing spectrum (Deloitte, 2017, p. 8). The convergence of 5G, IoT, and Edge Cloud computing is poised to generate huge amounts of decentralised data and, thus, increase already existing challenges and concerns in relation to security and protection of personal data and privacy. Last but not least, there are implications in terms of strategic digital autonomy and external dependency. Europe's choices on standards are not irrelevant in this respect. According to Albrycht & Swiatkowska: "*the integrity of the EU's 5G networks will rely on avoiding and minimising dependencies which threaten to result in 5G security breaches by third countries, especially those that are not like-minded*" (2019, p. 3). This argument can be applied by extension also to IoT and cloud computing and is transparently made in the presentation of the European cloud Gaia-X project by the German Ministry of Economy (BMWi, 2019).

Discussing the IoT, Deloitte calls governments to treat it as they did in the past when they were the main builders and providers of basic infrastructures (2017, pp. 5-8); a similar call can be justified for 5G and future cloud computing. There are several arguments in favour of adopting a public utility regime for these digital infrastructures. First, tactically, since they are still emerging such regime may be less complex and controversial compared to online platforms and AI, where vested interests and already-established business practices are a source of political resistance and technical complexity. Second, for 5G the current capital investment bottleneck is an opportunity for policy makers to support deployment and, consequently, lead and steer the process imposing regulatory requirements. To face this challenge, in fact, policy-makers "*can use a range of legal and regulatory actions to facilitate 5G network deployment. These include supporting the use of affordable wireless coverage (e.g. through sub-1 GHz bands) to reduce the digital divide, commercial incentives such as grants, or PPPs to stimulate investment in 5G networks*" (ITU, 2019, p. 68). Taking this direction would then give levers for regulation. Third, governments can provide good examples and use cases in the use of these infrastructures for the performance of various functions. By doing so, they can clarify business models and also establish good practices for transparency, security, and personal data protection. Fourth, governments' active role in developing standards can be justified on grounds both of speeding up deployment/innovation and of preserving strategic digital autonomy and ensuring cyber-security. Fifth, without governments' intervention as both regulator and infrastructure provider to ensure effective and equitable allocation of scarce bandwidth, excessive competition for spectrum would substantially slow down deployment. Last but not least, governments both as regulator and as user are best positioned to build trust and confidence in new emerging connected technology with respect to security and personal data protection, avoiding the backlash that currently concerns especially online platforms and AI. Actually, one may optimistically expect that government-led good practices and good governance of these new emerging technologies with regard to security and personal data protection may eventually produce positive overspills also on online platforms and AI.

A final point to be made in his section, which is not directly related to the discussion on public utilities, concerns the very controversial domain of competition law potentially applicable to online platforms. Already in 2017 it was argued that tech giants are posing a threat not so much because of their size but as a result of the enormous power they derive from controlling the data, which changes the nature of competition (The Economist, 2017b). They can anticipate trends and, thus, acquire new companies that may disrupt them, as in the case of Facebook's $22bn purchase of WhatsApp in 2014 seeming a 'shoot-out acquisition'. At that time The Economist proposed various new measures not falling into traditional anti-trust intervention, such as: considering companies' data assets when assessing merger requests and the price as signal of incumbent buying an emerging threat, identifying colluding algorithms, and giving more control on data to those supplying them. Then, as mentioned in 3.1.1 above, in 2019 news came that a Democratic presidential candidate proposes to unwind what are seen as anticompetitive tech mergers. Competition regulation in the context of the digital transformation and with specific respect to online platforms requires a radical renewal of a regulatory regime that was developed for the industrial era and as such is no longer appropriate or useful (Cohen, 2016). It needs regulatory innovation and the underlying rethinking and refinement of key concepts (i.e., market power) that has started only recently, mostly in Europe and not yet in the US.

3.1.3 DATA GOVERNANCE

It should be remembered here, as explained in Chapter 1, that in this study we focus on personal data. We can distinguish three approaches.

First, it is possible to envisage interventions going in the direction of decentralisation and individuals' data sovereignty with citizens enabled (if willing) to take control of their personal data. This would entail a number of other related steps, such as support for the deployment of platforms or application ecosystems based on personal data-stores and introduction of expiration dates for exclusive access to some data assets (in a fashion similar to copyright expiration). These could differ for personal vs non-personal data and could vary by sector. They could ensure that companies can keep engaging in data-driven innovation with a lower entry barrier (in terms of access to initial data assets). Other elements could encourage cloud infrastructures for personal data-stores, foster agreements/standards on the structure of personal data stored on online social networks and other online platforms and support personal data portability across online platforms. This kind of intervention is in line with the position and proposals from innovative projects such as SOLID and the platform proposed by the Ocean Protocol Foundation (OPF). These projects aim to equalize the opportunity to access data, so that a much broader range of AI practitioners can create value from it, and in turn spread the power of data. We must also respect privacy needs, which implies we must include privacy-preserving computation. The OPF is developing a protocol and network – a tokenised ecosystem – that incentivises making AI data and services available. One of the key proposals to change this oligopolistic environment has been to let users take control of their personal data by keeping their data stored in personal online repositories to which they can provide third parties with access after the latter obtain meaningful consent on access and use. For instance, as proposed by Tim Berners-Lee and embedded in his new project SOLID[78].

Second, a proposal that in some way could be related to giving users control over their data is that of treating data as labour and creating a new market. Treating data as labour may countervail the monopsony power of the tech giants (and particularly Google and Facebook), making data more available for other companies that may train their data analytics system and unleash productivity gains (Arrieta-Ibarra et al., 2018)[79]. In this respect it is important to stress that implementation (i.e. having data to train machine learning systems) is now possibly more important than introducing new analytics innovation. The more data the more AI systems learn and the more they can yield productivity gain. This solution may also help offset current concerns about AI reducing employment and worsening income distribution[80], considering also that Google and Facebook have market capitalisation similar to traditional large companies but employ one to two orders of magnitude fewer workers (Arrieta-Ibarra et al., 2018, p. 38).

Third, the proposal that competition regulators start opening algorithms, as advocated in the The Economist article cited above (3.1.2), brings us to the issue of 'algorithmic governance'. This concerns AI and also the actual deployment and application of a few key articles of GDPR (see discussion provided, respectively, in Sections 2.1.5 and 2.2.4). There is the view that regulation of algorithms is and will remain technically impossible and that Article 22 of GDPR (right to explanation and more in general requiring algorithm transparency by law) is neither legally binding nor applicable in the future. As seen in 3.1.1 above, it is argued that regulators either innovate and open algorithmic black boxes, or they abdicate to their role. A pragmatic proposal is to introduce a requirement for providing counterfactual explanation of algorithmic decision to achieve the same objective of Article 22 without opening the black box and without imposing too much burden on industry players. Alternatively, regulators may decide to impose algorithmic transparency by law (i.e. full application of Article 22). This will require a patient and steady work of innovation in regulatory mechanisms and technical capabilities and of collaboration with industry and academia. It will also impose a burden on industry players and, as in the case of the rights to withdraw consent and to be forgotten, create serious problems to existing practices, and potentially limit further future development and application of AI.

3.1.4 CYBERSECURITY

On cybersecurity, Aggarwal & Reddie (2018a, pp. 6-8) review a number of possible industrial policy measures:

- **Market creation**: Markets are created through rights, incentives, and opportunities. The case of China is a textbook case since the government is the sole customer for cybersecurity products created by state-sponsored entities (Cheung 2018). In France, coordinated procurement to build national capacities has been introduced to boost the national industry (D'Elia 2018). In the United States, the government and military try to support the industry with government-linked venture capital (Aggarwal & Reddie, 2018b).

- **Market facilitation** policies try to improve the functioning of the market, reducing transaction costs or providing incentives (documented for US, China, Japan, and Finland in, respectively, Aggarwal & Reddie, 2018b; Cheung, 2018; Bartlett, 2018; Griffith, 2018).

- **Market modification** through use of regulations to change the conduct of subjects. In France, for instance, the attempt to create a voluntary information-sharing (CERT-FR) that provides a reporting mechanism and shares best practices among companies, government sponsorship of crisis management exercises (D'Elia 2018). Followed at EU level in new measures (Timmers 2018), and with the new Cybersecurity Act of 2019 introducing an EU wide certification.

▪ **Market proscription** involving government measures that attempt to prohibit specific behaviours. Export controls emerging at EU level (Timmers 2018) and procurement rules (most typical of France, D'Elia, 2018) are perhaps the most obvious examples.

Out of these, market proscription measures may not be needed, whereas market modification measures that deepen what is already foreseen in the EU GDPR and in the Cybersecurity Act seem feasible and desirable. For instance: provide common ground and advice for avoiding cybersecurity breach incidents; enforce the certification methods in order to convince organisations to improve their security measures; impose heavy financial penalties comparable to cybersecurity breach incidents' total cost on victim organisation.

We can conclude this overview of possible policy and regulatory responses highlighting that the right balance will probably be struck somewhere in the middle, between the two extremes of 'leaving it to the market' and 'make it a utility'.

3.2 Proposed scenarios and their high-level assessment

In this section we present the proposed scenarios constructed along two dimensions: a) digital infrastructures regulation ranging from 'hands off' to 'interventionist' (resonating the discussion of general approaches and of public utility debates presented in 3.1.1); b) personal data regulation ranging from weak (no control to data subjects) to strong (data subject sovereignty, so referring to individual and not state sovereignty). Before graphically presenting and then textually describing the four scenarios, some clarifications are in order for a correct reading.

First, methodologically, it is consolidated practice to make scenarios extreme, also through several simplifications. This enables to capture collectively (through all four scenarios) most of the possible features that will characterise the actual future, including those aspects policy makers may want to avoid. Scenarios are just means to the end of identifying implications for policy having in mind that the actual and/or desirable future will result from a combination of features from different scenarios. It is important to stress that each scenario may also contain elements one can see in the current situation, but they are part of more radical, extreme developments envisaged. So, if one scenario contains elements of the status quo the fact that this scenario is not desirable should not be automatically extended to features that reflect existing conditions. The latter must be seen as part of an extreme scenario and not judged undesirable as such but only as long as they contribute to the extremization of the scenario narrative.

Second, the scenarios here rest on the important simplification that they are generically European, without entering into the potential differences between, on the one hand EU level and Member State level, and on the other into the differences and peculiarities of 28 different states. This is an inevitable simplification to avoid presenting 29 (EU plus 28 Member states) different scenarios. Yet, we can anticipate right here that a unified European approach would be highly desirable and much more effective and efficient.

Third, as we anticipated in Chapter 1, when discussing 'digital infrastructures' we focus only on four that are exemplificative and not exhaustive, and we only briefly mention the issue of machine data with regard to data protection. There are at least two 'infrastructures' that we did not consider and could be associated with the data protection dimension: they are blockchain and eID as they could both enable (though in different ways) user sovereignty; they are mentioned in passing in the narrative description of the scenarios.

Finally, this study and certainly the final scenarios are a brave exercise in complexity reduction. As anticipated at the very start of this Chapter, the two dimensions of analysis used (data protection and digital infrastructures) are closely interlinked with implications possibly entailing a wide mix of policy and regulatory measures. As we hinted there, although we focus on personal data, the issue of machine data should be considered especially for mixed data sets (containing both personal and machine data) for current guidance on how GDPR and the FDD Regulation leave a blurred area that may lead to stricter regulation if the former would prevail on the latter. Furthermore, other dimensions may be needed but would make the scenarios less intelligible and less manageable. Both axes could be split into two sub-dimensions: 'what' (four infrastructures; several different ways to regulate data governance) and 'how' (hands off vs interventionist; weak vs strong) and further refined with a cyber security dimension: cybersecurity regulation (possibly 'hands-off' vs 'interventionist'). The same holds for data regulation. The report provides a simplified model to explore the possible scenarios and available policy options by striking a trade-off on the number of dimensions and on the nuances in each dimension. As part of that trade-off, some of these nuances are only briefly considered in the narrative description of the scenarios, and cybersecurity is de facto subsumed under the digital infrastructure regulation dimension. It is assumed here that an interventionist approach would entail also more stringent cybersecurity state-emanated measures (i.e., with higher sanctions), whereas under a 'hands-off' approach cybersecurity would rely on softer means while still having security as an important objective. The issue of machine data, which is considered out of scope of this study, is mentioned briefly below and again in Section 3.3 in relation to the policy implications of the scenarios.

An intuitive way to render this discussion graphically and to prepare the presentation and illustration of the scenarios is the stylised 'regulation equaliser' depicted in Figure 8 below. The vertical dimensions convey the idea that regulation may

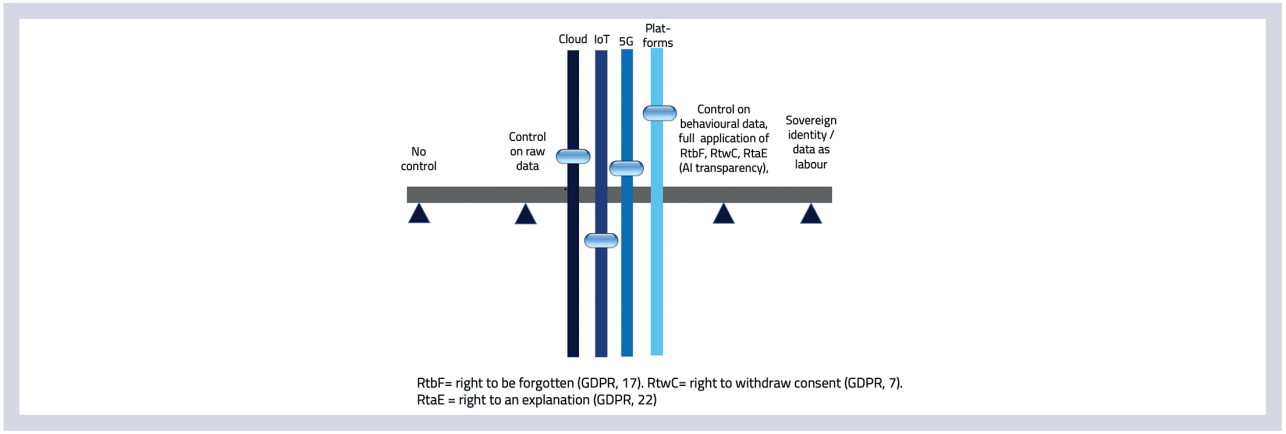


Figure 8: The Regulation Equaliser

address different infrastructures with different intensity and nuances. The horizontal axis indicates from weak to strong some specific actions to grant control over personal data to individuals. In the stronger side of data regulation, we envisage the full application of the GDPR rights to be forgotten, to withdraw consent, and to an explanation that would mean requiring in a binding way that algorithms are transparent. In this strongest approach we have also included the situation whereby for mixed datasets (including personal and machine data) GDPR would prevail as the main sources of regulation when the two categories of data are inextricably linked, even when the personal parts are very small compared to the machine parts.

With the above clarifications and disclaimers, Figure 9 presents the four scenarios we have identified. As is often the case with 2x2 scenarios matrices, it is immediately intuitive that some are more suitable than others to be the contender for the desired future. It is the case that some scenarios are very unlikely and/or less relevant for policy and regulation and/or clearly not desirable.

The scenarios are described below by considering the position of the key stakeholders, being governments, businesses, citizens and regulators, and analysed with respect to the following key objectives: economic growth, innovation, trust (i.e. from the user perspective), level playing field (i.e. from the supplier perspective) and fairness (i.e. equitable access to economic opportunity).

Ultra-Liberal (1): soft infrastructure control and weak data protection

In this scenario, there are few policies in place, regulators are a small player and governments take a hands-off approach to the infrastructure and there is a lot of freedom in handling data. Businesses that provide infrastructure and services have a lot of room to operate. Citizens will experience little protection and mainly influence by choosing what to use and buy, assuming choice exists. In this scenario business is the strong player.

Deployment of new infrastructure like 5G will be driven by market opportunities mainly, which may result in availability only in densely populated urban areas. Also, IoT and cloud development are left to the market, as well as industry development and standardisation/self-regulation efforts. Cybersecurity would be pursued through co-regulation, self-regulation, and standardisation rather than strict governmental regulation.

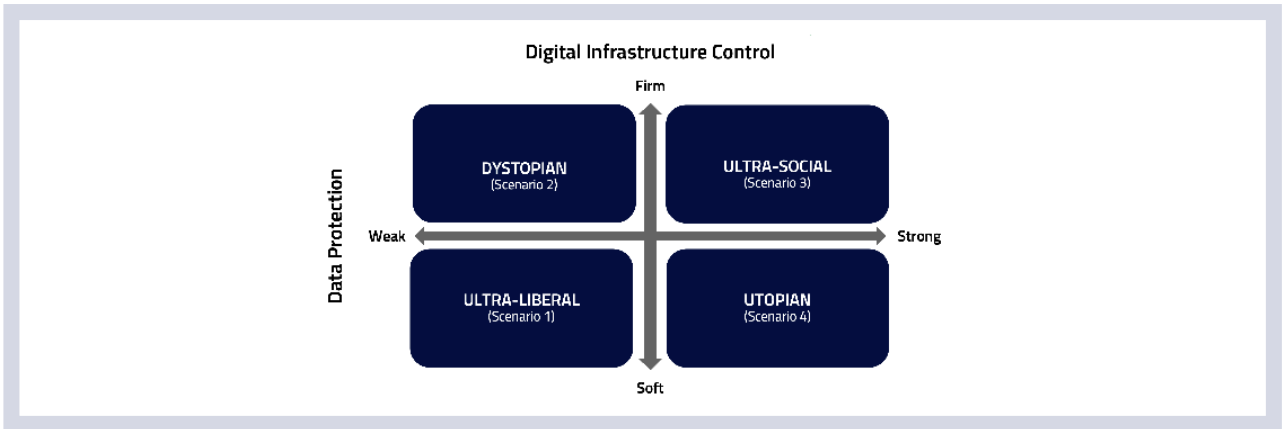


Figure 9: Scenarios

Data will to a large extent be controlled by large private enterprises, which continue to extract behavioural surplus without effective oversight and effective sanctions. Online platforms and tech giants can increase their advantages in terms of access to data which in turn enables continuous learning and improvement of their algorithms.

Economic growth in this scenario will be mainly business driven. Infrastructure investments have a multiplier effect on both short- and long-term economic growth. Without active government involvement in driving innovative infrastructure such as 5G, deployment might be delayed due to lack of short-term financial resources or returns. Innovation generally benefits from government stimulation, a hands-off government with respect to infrastructure may lead to less innovation in this scenario. Weak data protection has two sides when it comes to innovation, on the one hand it allows new ways of using data which fuels innovation, while on the other hand it may prevent citizens adopting new technology resulting from privacy concerns for example.

Trust for citizens comes down to trusting governments, business and regulators. Given that this scenario is mainly driven by business, it all depends on the trust citizens have in these businesses. Something that will vary from business to business. Since the government acts hands-off this scenario has a high likelihood of the winner takes it all. Since there is little regulation both on infrastructure and data, dominant market players will have ample possibilities to further strengthen their position. As a result, this scenario will highly unlikely produce a level playing field. Also, fairness is under pressure in this scenario, since the deployment of both infrastructure and services will be mainly market driven, as an example the deployment of 5G may be limited to densely populated urban areas, generating polarisation of access as thus no equitable access to economic opportunity (digital divide).

In this scenario neither Europe's technological sovereignty nor individual data sovereignty for European citizens are likely to emerge. Imbalances in the European data economy (export raw data, import refined results) will likely not be removed, and guidelines about data processing and ownership most likely remain without tangible results.

Dystopian (2): firm infrastructure control and weak data protection

In this scenario, governments control the infrastructure while there is freedom in handling data. The role of the regulator depends on the government approach of either strong regulatory control or public utility ownership. Businesses that provide infrastructure will face government intervention either directly or via strong government-controlled regulator. Businesses that provide data driven services have more room to operate but nevertheless can expect government interference due to the fact that the data travels over government-controlled infrastructures. Citizens will experience that access to, and use of, infrastructure and to a lesser extent services and platforms is directly or indirectly controlled by governments. In this scenario government is the strong player.

Deployment of new infrastructure like 5G, IoT and Cloud will be driven by governments taking into account economic development and geopolitical development. As a result, governments may choose to work closely with a limited set of trusted infrastructure providers. Infrastructure cybersecurity will be pursued through strict governmental regulation.

Lack of data protection and strong government control over the infrastructure also gives governments ample opportunities to control the data, either directly or through private enterprises. Online platforms and tech giants can only increase their advantages in terms of access to data with government (in)direct consent. Imbalances in the European data economy (export raw data, import refined results) will likely not be removed.

Economic growth in this scenario will be mainly driven by government and selected businesses. Active government involvement in driving deployment of innovative infrastructure will boost global economic competitiveness. Innovation generally benefits from government stimulation, but at same time requires freedom for experimentation and alternatives. Too much government control may lead to less innovation in this scenario.

In this scenario, trust from a user perspective is mostly relying on trust in the government. Since the government and government-selected businesses are dominant this scenario will not produce a level playing field. Fairness is determined by the government in this scenario, since the deployment of both infrastructure and services will be mainly government driven.

This scenario with strong government intervention and without personal data protection is considered so much inconsistent with the European values, that it is not a viable option for Europe.

Ultra-Social (3): firm infrastructure control and strong data protection

This scenario combines governments control over the infrastructure with strong data protection. Given that regulation originates from parliament in democracies and is controlled by the regulator, the regulator plays an important role in this scenario with strong data protection regulation. It is also likely in this scenario that governments exert their infrastructure control mostly via strong regulatory control rather than full public utility ownership, which further strengthens the role of the regulator. Businesses that provide infrastructure will face government intervention via an

empowered regulator. Businesses that provide data driven services can expect regulator interference. Citizens will experience a mix of government control and regulator interference, where the regulator safeguards data protection of citizens also towards governments. In this scenario the regulator is the strong player.

With respect to the digital infrastructures, the combination of firm government control and strong data protection could lead to public private partnerships for higher level infrastructure layers like clouds and platforms. The public sector would invest to overcome the barrier of high capex and to safeguard inclusion. Examples might be health or education Cloud platforms deployed on top of a combination of fixed, 5G and IoT networks. To encourage private co-vestment, the policy chosen could call for a lower intensity regulation. Such a policy could nonetheless include new rules and decisions on digital competition policy (monitoring of anti-competitive mergers, considering price and data assets, new definition of market power, auditing collusive algorithms, etc.). In addition to direct regulatory action, the government as a user and provider of digital infrastructures could establish good practices in data exploitation.

In this scenario data protection regulations such as GDPR should be fully implemented and new measures and policy actions for individual data (both raw and behavioural), rights to be forgotten, to withdraw consent, and to explanation, with algorithm transparency being mandatory and legally binding. An identity system would be guaranteed by law and regulation and made possible through the adoption of (sovereign) eID solutions.

Economic growth in this scenario will be mainly driven by public-private partnerships. Active government involvement in driving deployment of innovative infrastructure will boost global economic competitiveness. Innovation benefits from government stimulation and public private partnerships. Strong data protection has again two sides when it comes to innovation, on the one hand it restricts new ways of using data which hinders innovation, while on the other hand it may take away concerns from citizens in adopting new technology.

In this scenario, trust from a user perspective is mostly relying on trust in the regulator. The independent position of regulators in democracies and the fact that regulators also protect citizens interests, should increase trust in digital. Since the government is firmly controlling the infrastructure it may not be a full level playing field. Regarding data-driven services the likelihood of a level playing field is higher given the regulator ability to safeguard data protection and act against dominance. Fairness is determined by the regulator and to a lesser extent by the government in this scenario.

In principle, this scenario would allow for strengthening both Europe's technological sovereignty and individual data sovereignty for European citizens in a world without regulatory frictions and unintended effects.

Utopian (4): soft infrastructure control and strong data protection

In this scenario, governments take a hands-off approach to the infrastructure, while there is a strong data protection. Businesses that provide infrastructure have a lot of room to operate. Citizens experience data protection and can influence success of business by choosing what to use and buy, while their interests are safeguarded by the regulator. In this scenario citizens are the strong player.

Infrastructure development and deployment will be market driven with governments staying at arm's length. A more open playing field for data-driven service providers may lead to accelerated infrastructure deployment due to larger demand and faster uptake of services by users.

Data protection will safeguard citizens interests and the combination of freedom to operate on the infrastructure side may turn out to be a fertile environment for the development and deployment of trusted data-driven services.

Economic growth in this scenario will be mainly driven by a combination of technology push by businesses and market pull by citizens willing to explore and use trusted services. Innovation in this scenario will be driven by ecosystems that bring together businesses, innovators, entrepreneurs and early adopter citizens, but the business environment will show uncertainties by lack of regulation.

In this scenario, trust from a user perspective is mostly relying on the combination of the regulator and the diversity of businesses due to the lack of dominant market players. Since the citizens, supported by the regulator, are the key actors in this scenario it is likely to produce a level playing field for data use, but not necessarily for commercially driven infrastructure. Fairness is determined by the regulator in this scenario, and to a certain extent to the citizens preferences.

In this scenario individual data sovereignty for European citizens is likely to emerge. It is also possible to achieve Europe's technological sovereignty, although this will not come as long as digital infrastructures are not regulated and incumbent tech giants are left untouched, since regulatory intervention needs real levers in the absence of any form of regulation of digital infrastructures.

In the description of the scenarios we have used insights from available theoretical and speculative reasoning extracted from the sources reviewed in this study. They give a qualitative assessment of their likely impact on the five policy objectives. economic growth, innovation, trust (i.e. from the user perspective), level playing field (i.e. from the supplier perspective) and fairness (i.e. equitable access to economic opportunity).

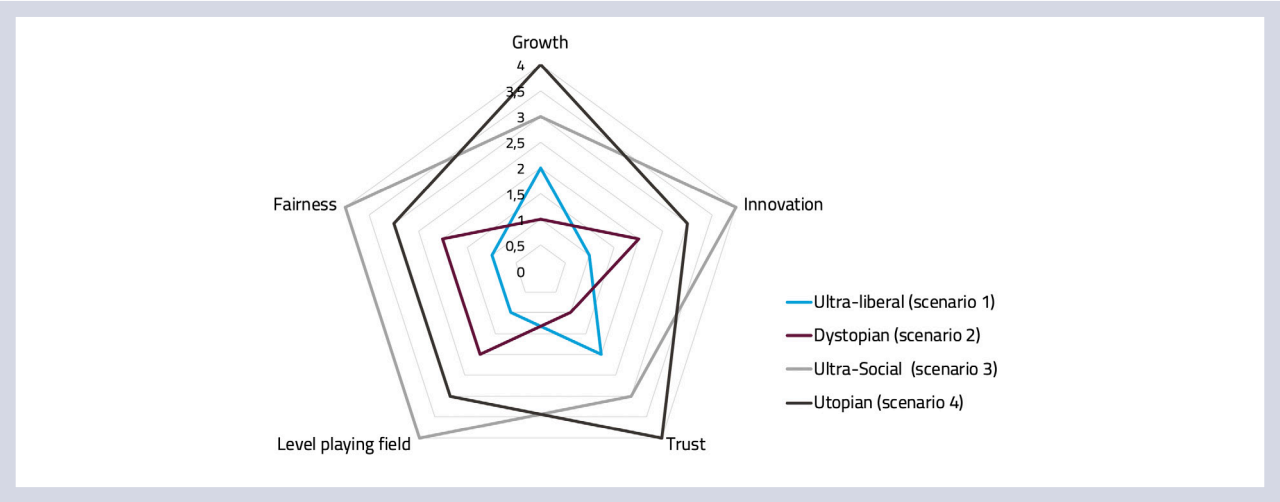


Figure 10: Radar assessment of scenarios impacts

Based on the qualitative assessments per scenario, for each policy objective the scenarios have been placed in the strict order from least (1) to most impact (4), thus providing a relative comparison between the scenarios. This has been depicted in the spider diagram below. Note that using a strict order forces a strong discrimination between the scenarios and leaves less room for nuances. The spider diagram therefore magnifies differences and has to be seen as a tool to give a quick insight into the relative strengths and weaknesses of the scenarios, rather than absolute differences.

The impact assessment shows that the ultra-social and utopian scenarios are very similar and deliver the most balanced overall result. The ultra-social scenario delivers somewhat better on fairness and level playing field due to the strong role of the regulator in that scenario. Also, the government hands-on attitude is assumed to be translated in relatively high public investments in research and innovation. Both Scenarios 1 and 2 suffer in particular of lack of fairness, level playing field and trust that has also negative effects on growth and innovation. Moreover, public Investments in infrastructure (missing in scenario 1) have a multiplier effect on both short- and long-term economic growth. In scenario 2 the state holds business in check, hampering also innovation. The investments in infrastructure have first order effects that should be stronger than the indirect effects from leaving the growth and innovation mostly to tech giants strengthening the status quo.

It is important to add some considerations for a correct reading of the diagram. First, our scenarios are not based on a combination of very well defined and already tested policy measures or regulations, as would be the case for instance for taxation with empirical evidence on potential effects available. The scenarios are a combination of general approaches to policy and regulation not anchored to specific and concrete measures for the simple reason that such measures still have to be formulated and/or assessed in terms of their costs and benefits. Hence, this creates an element of subjective judgement in the assessment that cannot be eliminated.

Secondly, these scenarios represent the more extreme choices, while in reality one finds mixed approaches that combine measures from different scenarios.

True innovation is spurred where more and more innovators will have access to data, and where a balanced regulatory environment gives certainty and stimulation to industry (innovators-regulators debate discussed extensively above). Hence, this is more likely to occur in Scenario 3 and to a lesser extent in Scenario 4 Again this is not to say that there will be no innovation in Scenario 1 and 2, but it will be less and more led by either government or by existing large tech giants. Fairness (i.e. equitable access to economic opportunity) and level playing field, it goes without saying, are lowest in Scenario 1 and highest in Scenario 3; Finally, with regard to trust it is important to explain that the high score for trust in Scenario 4 descends above all from full control over personal data and it is higher than in Scenario 3 as here it is dependent on trust in the regulators and government.

Given this assessment, the way forward would seems clear, since Scenario 3 is superior to the others and could simply take some elements from Scenario 1 to maximise also the innovation impact. Scenario 4 looks good in the idealistic extreme, but depends very much on the unregulated developments between the tech giants and the citizens/users. Yet, these are theoretical scenarios not considering all the difficulties of introducing perfect regulation producing the exact intended effects and not unintended and undesirable ones. As anticipated, only a mix of scenario features can identify a realistic approach offering a solution to the Regulators/Innovators dilemma presented in the Introduction (see Figure 1), to which we now turn in the next and final section.

3.3 Towards solutions and regulators/ innovators dilemma

The best way of picking up from the previous section is a quote from Juliet Cohen's recent book on the legal constructions of informational capitalism: "… Digital infrastructures are not simply instruments of innovation and liberation at the same time as law and regulation are not simply instruments for the promotion of just outcomes (arbiters of disputes or agents of modernisation). In different ways they both sit between truth and power" (Cohen, 2019, p. 4).

What truth and power refer to here is in a way the tension between values and economic interest or state interest and ideology, from which opposing claims and discourses emerge in the public debate. Neither 'leave it to the market' nor 'make it a public utility' are perfect in representing the full gamut of values, economic interests, and state priorities. Digital infrastructures if totally unregulated do not automatically ensure distributed innovation and equitable economic opportunity and growth. Furthermore, characterising the current context 'laissez-faire' or 'leave it to the market' would not foster innovation for it would leave uncertainty and anti-competitive positions intact. In the same way interventionist regulation on both digital infrastructures and data protection would not necessarily produce the desired outcome and may as well delay innovation if not well calibrated and implemented in a specific way.

Time and again well-designed policies have unintended effects in the implementation process when heterogeneous players with different objectives and behavioural strategies and biases enter into the picture. We might identify a possible mix considered more reasonable to have a less strict regulatory approach to machine data. This means, for instance, that for mixed datasets combining personal and machine data GDPR should not automatically prevail as the main source of regulation. More in general, machine data should be regulated and security issues be considered (i.e., avoid loss of trade secrets). But strict regulation on machine data should be avoided for it may hamper the full development of the European data economy and Industry 4.0 innovative potentials.

In a possible mix (but many others can be envisioned) digital infrastructures are regulated with different degrees of interventionism, whereas the data protection area of policy intervention is positioned somewhere central on the horizontal axis. Digital infrastructure regulatory interventions must be considered in different gradations within an integrated structure due to their interdependence; for example, interventions on cloud infrastructures can bear direct consequences on IoT infrastructures. However, regulation with regard to the sovereignty of individuals over their data can have a high degree of independence from infrastructure regulation; for example, individuals can have a high level of control over their data on tightly-regulated digital infrastructures or no control on completely unregulated ones. This could be one of the reasons for the introduction of policy interventions specifically on data regulation for citizens' sovereignty. Such interventions include the GDPR and the more recent directive on open data and the re-use of public sector information (2019/1024 of 20 June 2019), which makes explicit the requirement of personal data protection and of the use of personal data on the basis of consent of individuals or on a legal basis. A corollary is that this may also hold where future hard or soft law implements other fundamental rights such as for potential AI regulation [e.g. Art 1, 6, 7 and 21 of the Charter of Fundamental Rights of the European Union].

It can be observed that current policies in Europe are mainly a bit left of the centre in the scenario picture. Meanwhile, there is discussion of taking realistic and effective steps in the direction of mixing Scenario 3 and Scenario 4, gradually giving individuals sovereignty over their data, stepping up monitoring and action by competition authorities on online platforms and tech giants.

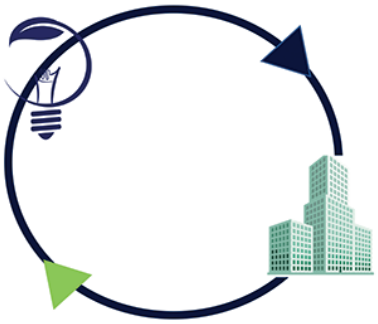
Next we discuss the 'how', providing the direction for the solution of the Regulators/ Innovators dilemma, as depicted in Figure 12.

Government acting as user, infrastructure provider, and as regulatory innovator in collaboration with the makers (innovators) can solve the dilemma and build the governance framework needed to spur innovation and build trust. Europe, through a firm coordinated action between the EU and the Member States, can virtuously connect makers (the innovators) and shapers (the regulators) in order to create an innovation enhancing governance and regulatory framework that respects European values and rights while creating economic opportunity for all users (individuals, companies or civil administrations). Regulatory innovation requires defining a mix of new mechanisms and capacities but also making political choices: adopt a precautionary approach when uncertainties concerning crucial and value-relevant issues require this and take a more stringent regulation. Or instead manage risks assessing the costs and benefits of regulation and, when the costs outweigh the benefits, use a softer approach or substitute by co-regulation, steering self-regulation, and collaborating with the innovators in the process of standardisation.

The objective of this report is not to propose concrete policy recommendations but rather to provide a roadmap on how policy makers can move toward making their selection on the regulator equaliser and implement it. Two general types of actions are needed: a) further scope possible policy options and analyse their consequences also with the help of future well focussed socio-economic and socio-technical research; b) building new capacity and mechanisms.

Many of the domains explored in this report remain terra incognita, unchartered areas where no ready and quick

Innovators
Know parameters, have tools, and can follow guidance for responsible innovation and better competitive conditions to disrupt incumbents.



Government acting as user, infrastructure provider, and regulatory innovator using the right mix of precautionary and risk management approach, guides secure, privacy compliant, and equitable developments.

Figure 11: The Regulators/Innovators solution, source: re-elaboration from Deloitte (2017, p. 3)

solutions are either available or have been tested and assessed before.

First, policy makers – possibly in collaboration with industry – should launch and fund a new stream of socio-economic and technical research aimed at better understanding the future development of new technologies and their potential socio-economic impacts and negative side effects under different levels of regulatory intervention. This would help map uncertainties and risks and decide when regulation could be decided simply on the basis of a cost-benefit analysis and when, instead, a precautionary approach would be needed.

Second, policy-makers should embark into institutional innovation to build new mechanisms, processes and capacity in order to find new ways to approach new phenomena, instead of using the old tool box for emergent developments that escape the reach of old ways of making policy and regulation. A few examples are briefly listed below:

- **Competition**. In a conference speech delivered on 9 December, 2019 Commissioner Vestager announced the intention to redefine the Commission Market Definition Note. On this aspect law scholars and economists have already produced insightful analysis and a proposal. A task force or high-level panel could quickly be set up to innovate the concepts and instruments of competition policy in the digital era. It is also urgent that regulators develop the internal capacity to scrutinize collusive algorithms and anti-competitive mergers looking also at the data implications rather than only at the old definition of market and market power.

- **Algorithms.** More research, dialogues with the machine learning community, and new internal capacity should be quickly pursued in order to decide how to best regulate algorithms with regard to their transparency and the need to avoid discriminatory decisions. It is important that policy makers enlist law scholars and AI scientists to move toward a technology neutral regulatory framework.

- **5G networks.** There are a lot of guess estimates and speculations about both the potential benefits and excessively high capital expenditure for the deployment of 5G networks. There is also an alleged dearth of use cases. Policy makers should push quickly for thorough socio-economic analysis of these issues as to produce the business case that would or would not justify the investments of public funds and the creation of new procurement and PPP mechanisms to ensure full and equitable deployment of 5G.

To sum up, more evidence and new mechanisms and capacities are needed for regulators to enter on a par into a dialogue with innovators so as to collaboratively solve the above described dilemma to rebuild trust and spur the development of an equitable and innovative data economy.

Yet, this is not to suggest a 'wait and see attitude' until more evidence is available. Actions can be taken: prioritising the areas of intervention, engaging with academia and industry in constructive dialogues on possible joint initiatives, pilots in prioritised areas, building capacity, innovating instruments and processes, and defining new streams of publicly-funded research on the most relevant topics.

4. TECHNICAL ANNEXES

4.1 Users: statistics, attitudes, behaviours and market failures

A Eurobarometer survey conducted in 2019 to explore awareness and attitudes about the GDPR one year after its introduction found that 67% of Europeans had heard about it but only 36% knew what it was (Eurobarometer, 2019, p. 3). Almost two-thirds (65%) had heard of the right to access their data, 61% had heard of the right to correct their data if it is wrong, 59% about the right to object to receiving direct marketing and 57% about the right to have their data deleted and forgotten. The three most exercised rights are the right to object to receiving direct marketing (24%), the right to access personal data (18%), and the right to correct personal data if it is wrong (16%). Just over one in five say they are always informed about the conditions attached to the collection and use of their personal data online, and only a minority (13%) fully read privacy statements online. The majority of social network users (56%) have tried to change the default privacy settings of their profile. The most common reason for not doing so are that users trust sites to set appropriate privacy settings (29%) or that they do not know how to do it (Eurobarometer, 2019, p.4).

Another Eurobarometer survey conducted in 2017 provides a picture on European attitudes and perspectives on cybersecurity (Eurobarometer, 2017). Over eight in ten (87%) see cybercrime as an important challenge, a significant increase on the 80% recorded in March 2015. The rise is even more significant when looking at the proportion of respondents who see cybercrime as a very important challenge: 56% compared with 42% in 2015. There are significant differences across countries in the proportions of respondents who think that cybercrime is a very important challenge, ranging from 76% in Cyprus, and 75% in the Netherlands to just 39% in Sweden and 26% in Estonia. Less than half (49%) of the respondents agree or mostly agree that law enforcement is doing enough to combat cybercrime, with the proportion of respondents who totally agree being generally low across Member States. Respondents express high levels of concern about the security of their online transactions; 73% of Internet users are concerned that their online personal information could not be kept secure by websites and 65% are concerned that their online personal information could not be kept secure by public authorities (Eurobarometer, 2017, p. 5). When asked to choose among a list of common risks when using the Internet, the two most common concerns mentioned by respondents are the misuse of personal data (45%) and the security of online payments (42%). Concerns about online privacy and security are having an impact on behaviour: among respondents that are Internet users, over six in ten (62%) have changed the access password of at least one online service during the last 12 months; 87% of respondents avoid disclosing personal information online; nearly half (45%) have installed or changed anti-virus software, and nearly four in ten (39%) have reduced the personal information they give out on websites. However, few have taken the step at reducing the goods and services they buy online (12%), opting out of conducting online transactions (11%) or opting out of online banking (10%). A majority of respondents are concerned about being the victims of various forms of cybercrime, with the largest proportions of respondents expressing concern about discovering malicious software on their device (69%), identity theft (69%) and bank card and online banking fraud (66%). Less than half of respondents have actually been a victim of the various forms of cybercrime. The two most common situations experienced by respondents are discovering malicious software on their device (42%) and receiving an email or phone call fraudulently asking for access to their computer, logins or personal details in 38% of cases (Eurobarometer, 2017, p. 6).

Additional insights on cybersecurity can be garnered from the results of the Eurostat household survey. According to Eurostat data, in the EU28 15% of respondents reported not buying online due to security concerns. This percentage is above 25% in countries such as France and Sweden and below 10% in Poland, Lithuania, Slovakia. The percentage of Europeans reporting having experienced security problems was 17% in 2015, down 5% point compare to 2010. With respect to this dimension there are sizeable country variations (i.e., 6% in the Netherlands versus 25% in France and 29% in Croatia). As per identification procedures the latest Eurostat data for 2018 show the following EU28 average (total does not sum to 100% because multiple answers were possible):

- Simple login with username and password as identification for online services: 70%;
- Social media login as identification for online services: 29%;
- Security token as identification for online services: 14%;
- Electronic certificate or card as identification for online services:15%;

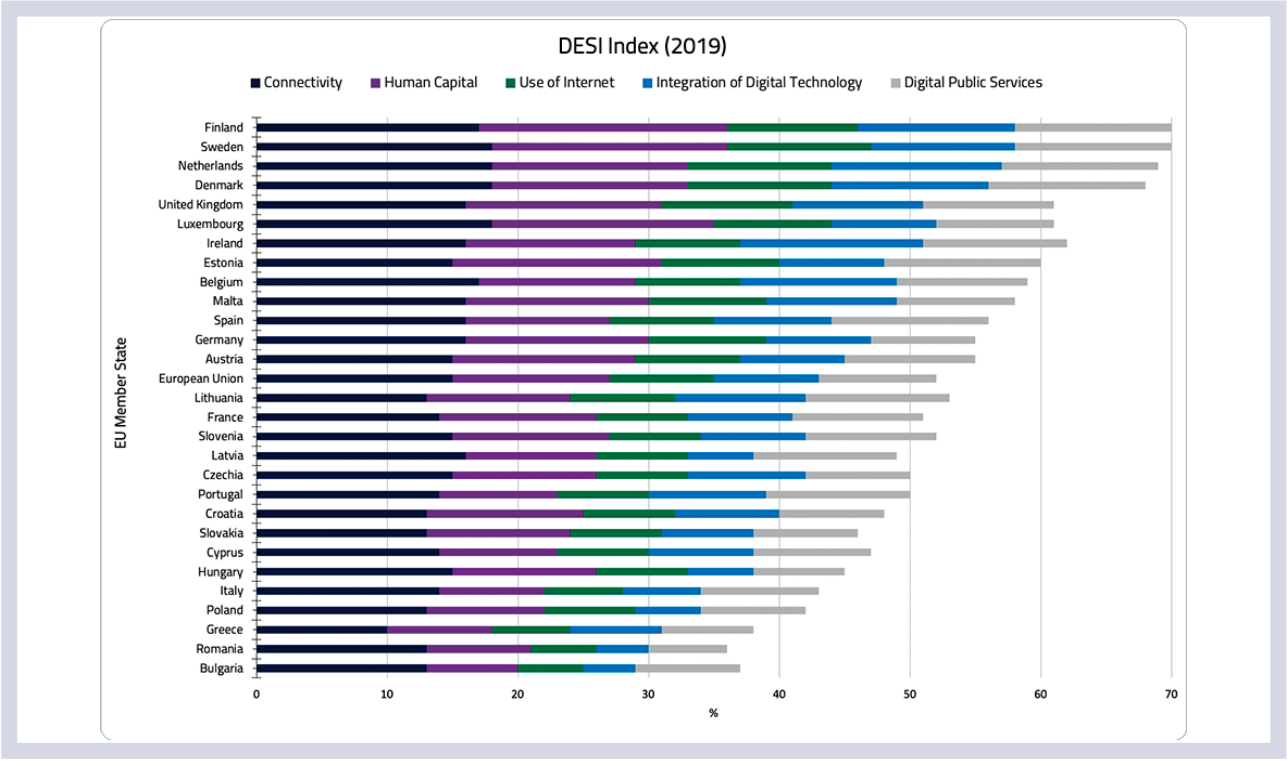


Figure 12: The situation in the EU regarding the 5 DESI indicators, source: https://digital-agenda-data.eu/datasets/desi/indicators

- Procedure involving their mobile phone: 38%;

- Single pin as identification for online services: 26%.

Cross-tabulation of these two variables against a measure of level of digital development in each country (measured by the European Commission Digital Economy and Society Index, DESI), are plotted below. The first picture (Figure 13) shows the digital landscape in the EU in 2019 as per the five DESI indicators (i.e. connectivity, human capital, integration of digital technology and digital public services), with Finland, Sweden, the Netherlands and Denmark having the highest index scores with respect to these indicators.

Figure 14 depicts the situation in 2015 matching the DESI indicators against the incidence of cybersecurity problems. EU Member States with the highest incidence of cybersecurity problems, such as Croatia, Portugal and Bulgaria, have a much lower DESI index score. This suggests that countries with better developed digital landscape may also have more robust cybersecurity programmes in place.

This assumption is supported by Figure 15, which shows the extent to which cybersecurity concerns impeded online purchasing behaviour in EU Member States in 2010 and 2015. In 2010 Spain, France, North Macedonia and Italy faced the highest effects on online purchasing behaviour. By 2015 North Macedonia still had the highest effect, but followed by Sweden, Romania and Portugal. Sweden being the only exception, these three figures show a clear correlation that countries with lower DESI index scores have higher cybersecurity risks. But how these risks appear (either in terms of compliance to cybersecurity measures or cybersecurity protocols being robust enough, along with regulatory pressures) will be further elaborated in later sections of the Technical Annexes..

In 2016 in EU28 59% of respondents reported knowing that cookies can be used to trace movements of people on the internet; the percentage is as low as 24% in Romania and as high as 84% in the Netherlands. In 2010 59% reported using any kind of IT security software or tool (anti-virus, anti-spam, firewall, etc.) in order to protect private computer and data but the same year only 15% reported always or almost always doing safety copies/back up files. Finally, some statistics on use of smart phones for EU28 based on Eurostat data for 2018 that are a proxy measure of how much European trust their devices:

- Individuals use a smartphone for private purposes: 75%

- Smartphone has some security system, installed automatically or provided with the operating system: 32%

- Smartphone has some security system, installed by somebody or subscribed to it: 12%

- Individuals already lost information, pictures, documents or other kind of data on the smartphone as a result of a virus or other hostile type of programs: 4%

- Individuals at least once restricted or refused access to

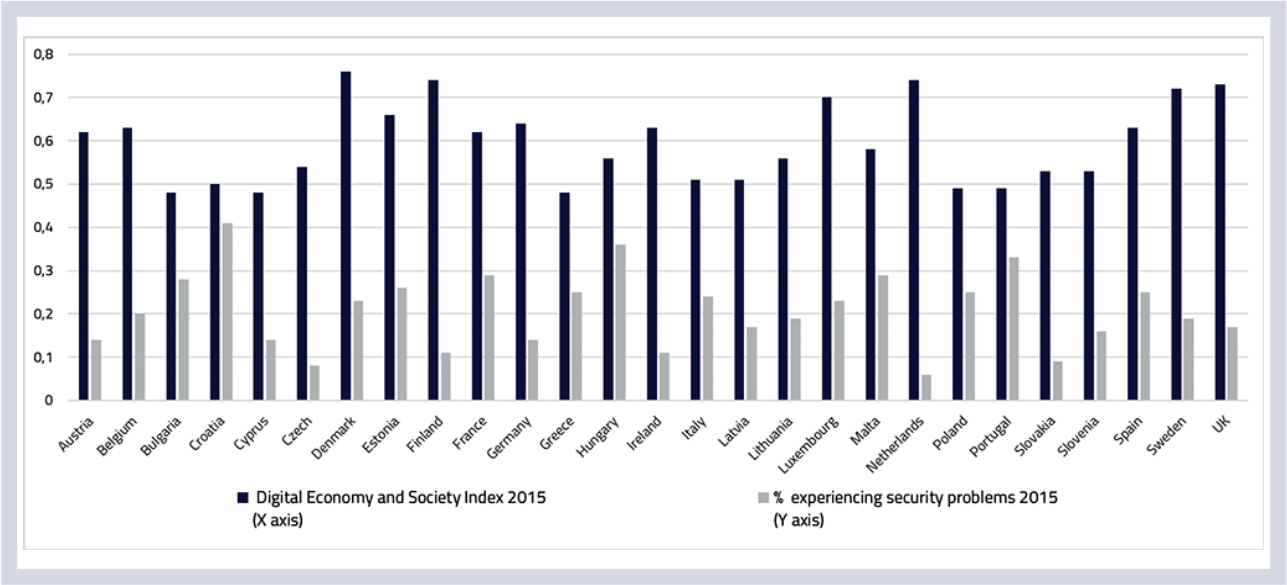


Figure 13: Cybersecurity problems and development of digital landscape, source: elaboration on https://digital-agenda-data.eu/datasets/desi/ indicators and https://ec.europa.eu/eurostat/data/database

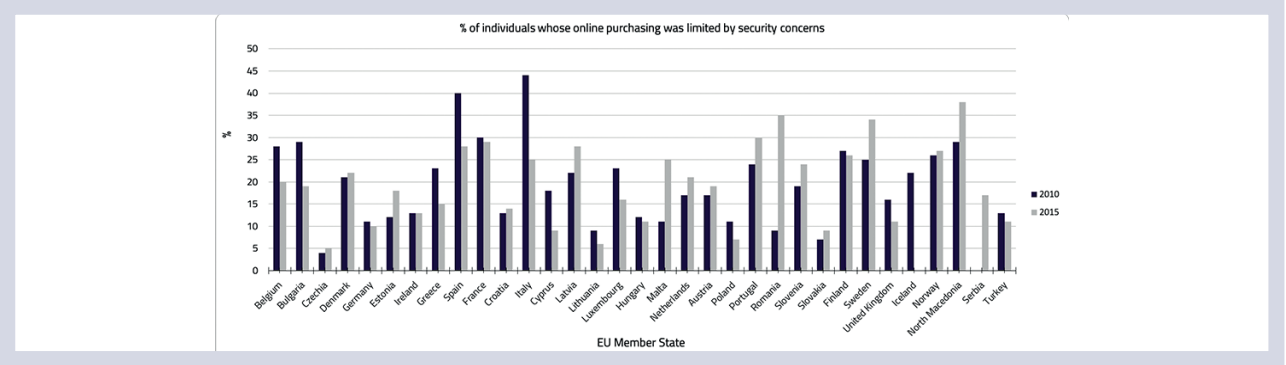


Figure 14: Cybersecurity problems and effects on online purchasing behaviour, source: https://ec.europa.eu/eurostat/data/database

personal data, when using or installing an app on the smartphone: 43%

▪ Individuals never restricted or refused access to personal data, when using or installing an app on the smartphone: 21%

▪ Individuals didn't know it was possible to restrict or refuse access to personal data, when using or installing an app on the smartphone: 5%.

There is no equivalent Eurobarometer on European firms and GDPR and the picture on this issue is preliminary and is based for the sake of illustration on a Europe wide survey conducted by a consulting company (RSM, 2019) and on a larger security breaches survey conducted by UK Department for Digital, Culture, Media and Sport but focussing only on UK companies (UK DDCMS, 2019). According to the first source (survey was conducted with about 600 firms between April and June 2019), one year after the introduction of GDPR 30% of European firms are still not compliant with its requirements. On the positive side it can be noted, however, that: 73% report that GDPR has led to improvement in management of customers data; 58% report that GDPR has encouraged new, innovative use of data; and 31% that GDPR has made their business more operationally effective. On the other hand, firms reported also that GDPR has: required increased investments in cybersecurity (62%), cost of compliance slowed growth (37%), and made difficult to work with non-European businesses (28%). The UK 2019 cybersecurity breaches confirm that GDPR has led firms to invest more in cybersecurity: firms that invested more in cybersecurity reported as reason GDPR requirements (68%) and holding increasing amounts of customers' data (55%). When asked if they agree with the statement 'GDPR compliance has resulted in more investment in cybersecurity' in 62% of cases firms answered affirmatively. Finally, 26% of UK firms do not yet consider themselves to be 100% compliant with GDPR.

The picture on EU firms and cybersecurity is updated only to 2015 and key data are presented in an ad hoc report by Eurostat (Eurostat, 2015), whose key findings are: almost one out of three enterprises in the EU28 had a formally-defined

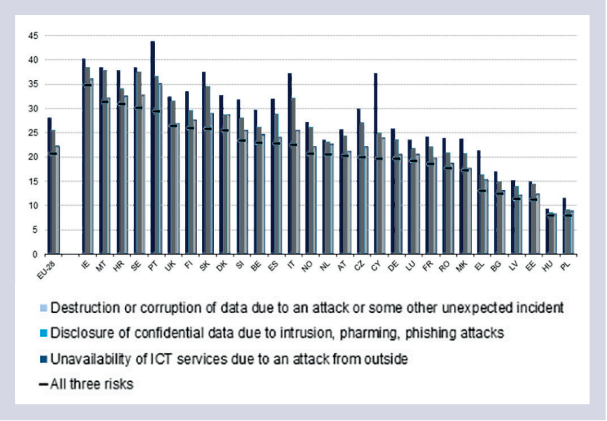


Figure 15: Firms addressing certain cyber risk by country (EU28, 2015), source: (Eurostat, 2015)

ICT security policy; the share of large enterprises that had a formally-defined ICT security policy was almost three times the share of small ones; the majority of enterprises having an ICT security policy (32% of the total), defined or reviewed their policy within the last 12 months (20% of the total); in all countries, most of the enterprises addressed the risk of destruction or corruption of data due to an attack or some other unexpected incident. Figure 16 shows the types of risks addressed by firms in different EU countries.

Figures 17 and 18 break down the group of firms with a formal cybersecurity policy in place, first by class size and then by sector of activity.

In the first graph the most noteworthy aspect was already highlighted above: there is a marked difference related to firms' size. When considering the breakdown in terms of sector of activity, as expected it emerges that ICT firms are considerably more active than firms in other sectors.

The data reported so far provide an interesting and informative picture and uncover some inconsistency between reported attitudes/concerns and actual behaviours (especially for individuals) that demand some additional considerations, borrowing from the perspectives of both the economics of

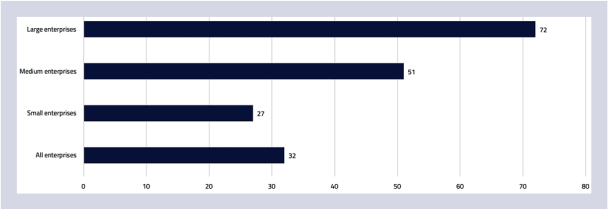


Figure 17: Firms having a formally defined cybersecurity policy by class size (EU28, 2015), source: (Eurostat, 2015)

privacy and information and from behavioural economics. These considerations enable to point out some potential market failures justifying regulatory interventions.

The leading behavioural scholars of privacy issues and other colleagues have amply demonstrated how it is too difficult for individuals to make reasoned and rational decisions on their personal data and the related expression of consent, due to behavioural bias and the complexity of the issues at stake and how they are rendered in terms of services and other contractual documents (Acquisti et al., 2015). Firstly, there is overwhelming evidence that most people neither read nor understand online privacy policies which users must accept before accessing digital services. Secondly, people struggle to make informed decisions about their informational privacy due to problems of bounded rationality and problems of aggregation. Thirdly, individuals' privacy preferences are highly malleable and context dependent. An impressive array of empirical privacy studies demonstrate that people experience considerable uncertainty about the importance of privacy owing to difficulties in ascertaining the potential consequences of privacy behaviour, often exacerbated by the intangible nature of many privacy harms (e.g., how harmful is it if a stranger becomes aware of one's life history?) and given that privacy is rarely an unalloyed good but typically involves trade-offs (Acquisti et al., 2015). Empirical studies demonstrate that individuals' privacy behaviours are easily influenced through environmental cues, such as defaults, and the design of web environments owing to pervasive reliance on heuristics and social norms. Because people are often 'at sea' when it comes to the consequences of their feelings

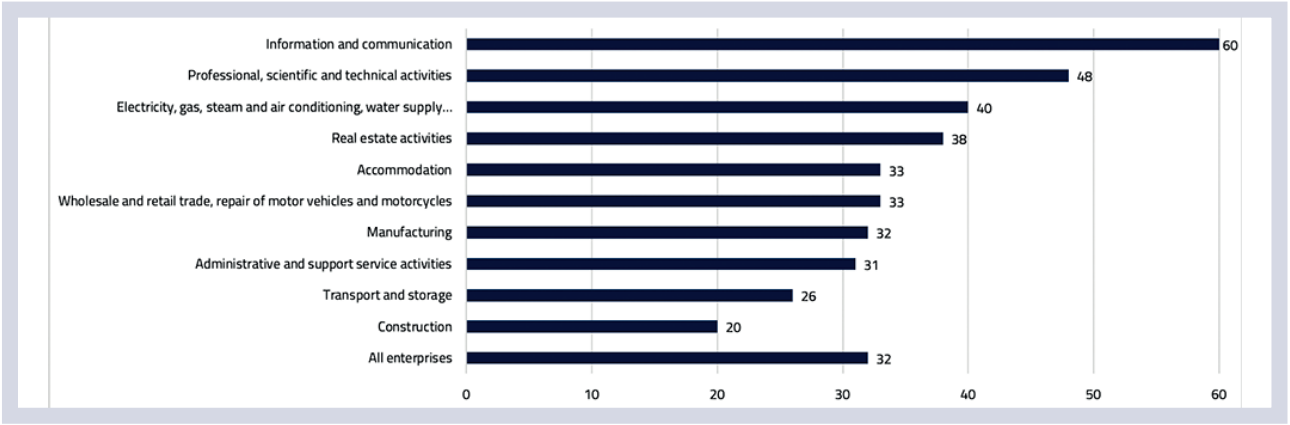


Figure 16: Firms with a formally defined cybersecurity policy by sector (EU28, 2015), source: (Eurostat, 2015)

about privacy, they typically cast around for cues in their environment to guide their behaviour, including the behaviour of others and their past experiences, so that one's privacy preferences are highly context dependent rather than stable and generalisable to a wide range of settings (Acquisti et al., 2015). According to Acquisti and his colleagues, this extensive uncertainty and context dependence imply that people cannot be counted on to navigate the complex trade-offs involving privacy in a self-interested fashion (Acquisti et al., 2015). Thus, many information law scholars seriously doubt that individual acceptance of the 'terms and conditions' offered by digital service providers (including Google, Facebook, Twitter and Amazon), typically indicated by clicking on a web page link, constitutes meaningful waiver of one's underlying rights to informational privacy (Solove, 2013, pp. 1880–1903). Legal scholars treat these agreements as 'contracts of adhesion' because they impose take-it-or-leave-it conditions on users who stick to them whether they like it or not (Zuboff, 2019, chap 2, Section V). In determining the potential loss from a breach of personal data, individuals may fall in various behavioural biases such as illusion of control ('I have done enough'), overconfidence bias ('yes it is a problem but will not happen to me'), and present bias (discounting future risks in exchange for immediate gratification).

Yet, even the most rational individual would have difficulty in avoiding present bias when making decision about personal data, due to some very specific aspects characterising consumer-to-seller information flows (Zhe Jin & Stivers, 2017). Data security and privacy policies are largely credence characteristics even with direct partners in a transaction. In addition, data persistence means that consumer valuation of an information flow is a function of the network of entities that access and use that flow. The complexity of this network, combined with the difficulty in credibly conveying and committing to these policies, creates an information problem: it is difficult for consumers to be fully informed of the potential network of decisions and outcomes, process that information, and decide whether to allow their private information to flow to the network. The opaqueness of the network often makes it difficult to establish causal linkages between seller policies and their effects, both positive and negative, on the consumer. In some cases, there may only be a probabilistic linkage between action and realised harm. This creates positive and negative externalities that no one actor may have full ability or incentive to internalise. The persistence of information means that transfer of private information is a sunk investment which could allow ex post (perhaps unilateral) renegotiation of how that information could be used or protected. Even an actor that knows all of the current protections and uses for their data cannot count on future use and protection. The commitment problem is especially relevant in privacy and data security, where perceptions about the value and use of consumer data is rapidly changing, and the technologies needed to control that data are also evolving. All of this creates a clear information asymmetry, meaning that individuals' decisions are potentially dependent on prior beliefs or assumptions about data policies and the trustworthiness of seller claims about those policies. Information asymmetry is one major market failure demanding regulatory intervention.

Moving to consider the perspective of the firms, it is evident that also for an executive deciding on an investment for protecting customers' personal data or for increased cybersecurity would be difficult. Lack of information, the steadily changing costs, and the same behavioural biases seen earlier may hinder a complete appraisal of costs and benefits to determine the ROI of the investment. Then procrastination and responsibility dumping may set in and the investment would be avoided. Even if the investment is made, then implementation may be ineffective when humans must perform tasks (e.g. enabling automatic updates, rebooting machines to apply some of those updates, or enrolling in two-factor authentication) and fail to do so, due mostly to present bias (Frik et al., 2018). Finally, because the indirect intangible costs of a security breach (related to loss of reputation potentially reverberating into losses in stock markets) are possibly higher than the direct costs, firms falling victims of a security breach may avoid publicly reporting it.

But security breaches do not only generate costs at firms which are directly affected. Interdependence between information systems allows breaches to propagate and negatively affect others (Kunreuther & Heal, 2003). In the language of public economics, a lack of firms' information security (for any of the reasons above, and either at the level of adoption of cybersecurity measures or at that of publicly reporting breaches) causes negative externalities in an economy. The presence of negative externalities justifies government intervention, for instance, in the form of laws aiming at reducing the costs of insecurity to society (Hiller & Russell, 2013).

4.2 Data protection and cybersecurity regulatory overview

4.2.1 EUROPE

With respect to Europe, as a response to the alarmingly fast rising (in numbers and severity) cybercrime and privacy breach incidents, cybersecurity and citizen privacy protection have become two of the EU's top priorities. One of the most important steps towards this direction was the establishment of ENISA. The European Union Agency for Network and Information Security (ENISA) is a centre of expertise for cybersecurity in Europe, which was established in 2004 by the EU Regulation No 460/2004. The main activities of ENISA include the following:

- Security/privacy related recommendations to interested organisations and companies.

- Support for security and privacy related policy making and

policy implementation.

▪ Direct collaboration with EU designated expert groups working in cybersecurity.

Then in the EU we have also the NIS Directive (Directive on Security of Network and Information Systems – EU 2016/1148) which was adopted by the European Parliament on 6 July 2016 and put into force in August 2016[83]. The NIS Directive's main goal is to achieve a harmonized framework of network and ICT security across the EU, as follows:

1. By improving cybersecurity defence levels at a national level.

2. By increasing collaboration and information sharing related to cybersecurity among all EU member countries.

3. By introducing specific cybersecurity mechanisms and incident reporting frameworks for Operators of Essential Services (OES) in Critical National Infrastructures (CNI) and Digital Service Providers (DSP) operating in the EU regardless of their originating country.

In the context of the NIS Directive, the EU member states are required to define their own regulations for financial penalties, similar to the GDPR, when cybersecurity breaches occur. Moreover, member states are required to take appropriate measures to prevent such violations. The NIS Directive's goal is to develop high cybersecurity levels across various industrial sectors that rely to a large extend on ICTs while, at the same time, they provide services and operations crucial for the support of economy and citizens' daily lives (e.g. utilities companies). These sectors, which are within the scope of the NIS Directive, are the following:

▪ Transportation (all means)

▪ Energy (Electricity, Oil and Gas)

▪ Banking (Credit institutions)

▪ Financial markets (Trading venues)

▪ Health (Health care providers and health system)

▪ Water (Drinking water suppliers and water distributors)

▪ Digital infrastructures (specifically, Domain Name Service (DNS) providers, Internet Exchange Points (IXP) operators as well as Top Level Domain (TLD) name registries).

▪ Digital Platform Services (DPS) (specifically the categories Cloud computing platforms and services; Search engines and Online markets).

In order to achieve compliance, Article 44 of the directive mandates that a culture of risk management, involving risk assessment and the implementation of security measures appropriate to the identified risks, should be supported and implemented by member states with respect to the protection of critical infrastructures. Furthermore, incident reporting obligations include any incident that affects the security of the ICT infrastructure such as electricity failures, hardware failures and cyberattacks of any kind and level of severity. Also, Article 19 of the NIS Directive promotes the deployment of European or other internationally accepted standards and mechanisms related to ICT security. In this context, the directive recommends the award of certifications in the two most widely accepted international standards on ICT security and business continuity management:

▪ ISO 27001: a best practice standard for ICT security.

▪ ISO 22301: an internationally accepted best practice standard on Information Security Management Systems (ISMS), which forms the basis of intelligent cybersecurity risk management and business continuity strategies.

Also, in Europe we had the following two recent breakthrough milestones in the European data protection and cybersecurity regulatory landscape:

▪ With respect to privacy, after four years of preparations, debates, and discussions the EU General Data Protection Regulation (GDPR) was approved by the EU Parliament on 14 April 2016[84] and was enforced on 25 May 2018 (the official website of GDPR is https://eugdpr.org/). It is considered one of the most important advances in data privacy regulations in 20 years.

▪ With respect to cybersecurity, on 7 June 2019, Regulation (EU) 2019/881 on ENISA (the European Union Agency for Network and Information Security or the European Union Agency for Cybersecurity, as it is named in the regulation document) and on the Information and communications technology cybersecurity certification, also known as the Cybersecurity Act, was published. The Cybersecurity Act[85] was enforced on 27 June 2019.

With respect to the GDPR, it replaced the rather outdated Data Protection Directive 95/46/EC and its main goals are the following:

▪ To harmonize data privacy legislation across European countries and help protect the privacy of European citizens.

▪ To raise privacy awareness among EU citizens and give them control of their data with respect to how they are used by private and public organisations.

▪ To ignite the development of improved EU-third country data transfer and data handling agreements and regulatory frameworks for the protection of EU-citizens' privacy.

▪ To motivate (and enforce, if necessary, through heavy financial penalties) organisations to rethink the way they approach and secure people's privacy.

There is huge discussion, controversy and exchange of views on an international level still, after more than one year of GDPR's enactment, that we cannot repeat it here. For our purposes, we should keep in mind the following points, with respect to regulations:

▪ Since the enactment of the GDPR, several cases of EU citizens' data mishandling from (mainly) US organisations and companies have been brought to court with favourable decisions (e.g. Google and Facebook cases) for EU citizens.

▪ The GDPR raised a privacy awareness wave across Europe whose repercussions and impact are felt also on other countries' efforts to set up privacy related regulatory frameworks.

▪ Through the financial penalties imposed by the GDPR on non-compliant organisations, it is expected that the privacy of European citizens will be better secured in the near future.

▪ Finally, the GDPR is expected to help bridge the gap between the different privacy cultures in different nations, especially the ones of the US and Europe. This will foster economic development and reduce the negative impact of cybersecurity breach incidents.

With respect to the Cybersecurity Act, the official confirmation of the upgraded pivotal role that ENISA assumes in fighting cybercrime across Europe within the Act must be noted. More specifically, Article 4, Objective 1 of the Cybersecurity Act clearly defines the new upgraded status of ENISA in the European cybersecurity landscape by stating that: "*ENISA shall be a centre of expertise on cybersecurity by virtue of its independence, the scientific and technical quality of the advice and assistance it delivers, the information it provides, the transparency of its operating procedures, the methods of operation, and its diligence in carrying out its tasks.*"

4.2.2 USA

The situation in the United States with respect to cybersecurity and privacy protection regulations has attracted worldwide attention. We believe this is due to the following reasons: the federated political system of governance; GAFA(M) which stands for Google, Amazon, Facebook, Apple (and occasionally Microsoft), also known as Big Four (or Big Five, with Microsoft); the cultural differences between US and Europe with respect to protection and use of people's (especially Europeans') personal data. In what follows, we will expand on these four points. Point 1 is, most probably, the reason for the rather unexpected fact that there is no general cybersecurity regulation available in the US. However, the states take measures separately such as by securing funds for improved security mechanisms as well as by demanding from governmental organisations or private businesses to implement cybersecurity measures, imposing fines in cases of cybersecurity breach incidents. In 2018, as reported by the National Conference of State Legislatures (NCSL), at least 35 states, D.C. and Puerto Rico have either introduced or considered more than 265 bills or measures related to cybersecurity. The key targets of the states' legislative initiatives include the following:

1. Inducing improvements over government cybersecurity practices.

2. Securing funding for cybersecurity R&D programmes and regulation initiatives.

3. Prohibit or reduce incidents of public disclosure of sensitive governmental cybersecurity related information.

4. Promoting cybersecurity related employment, training, and economic growth.

According to NCSL, at least 22 states had enacted 52 bills by the year 2018. NCSL's site lists these states along with the introduced bills, providing some description of their contents and mandates. Many of them refer to data collection and processing good practices along with the obligation of organisations to notify people soon after their data is compromised by a privacy breach incident. Business is subject to the specific cybersecurity legislation enacted in states and certain states have very stringent cybersecurity requirements for organisations, such as New York's regulations focused on the financial sector. One of the criticisms on these regulations, in contrast with the European GDPR, is that there are no clear penalties for organisations that fail to comply. California will put into effect its data privacy legislation in January 2020, with a goal to empower people to have more control over their personal information collected by organisations, similarly with the GDPR.

Going through the list of states that have taken the initiative to introduce a state cybersecurity legislation reveals a lack of sufficient common grounds on which to base a federal cybersecurity regulation effort, even though all states' legislation acts have reasonable requirements with respect to cybersecurity. The lack of a federal cybersecurity legislation, however, does not imply that businesses are not bound by rules. Government contractors must obey certain such rules. For instance, as of 31 December 2017 all contractors with the Department of Defence (DoD) must obey certain cybersecurity compliance requirements, otherwise the contract is terminated. Moreover, in January 2018, the General Services Administration (GSA) announced new regulations for contractors with respect, also, to handling data in an appropriate way and reporting privacy breach incidents as soon as they occur.

Outside of the federal U.S government, some industries also

have rules for data handling. For instance, health care is one sector governed by federal regulations for managing patient data. HIPAA, which stands for Health Insurance Portability and Accountability Act of 1996, is a US legislation act that mandates patients' data privacy rules and security provisions for protecting medical information (Title II, HIPAA Privacy and HIPAA Security rules).

In summary, with respect to point 1, the situation with respect to cybersecurity regulations appears fragmented and responsibility for defining and implementing such regulations is delegated to the states themselves, the various governmental agencies, and the different industries.

With respect to point 2, five of the biggest and most influential companies in the world have their central headquarters in the US. The complication with respect to cybersecurity arises due to the fact that these five companies store data and personal information of billions of people, a vast number of which reside outside US in Europe. Facebook and Google have already been in dispute, with the Court of Justice of the European Union (CJEU), with respect to their policies of doing business in the EU territory, part of which is their ways of collecting and storing EU citizens' personal information, contrary to what the GDPR mandates. Due to this fact, there have been numerous efforts by the US, jointly with the EU on several occasions, to bring about an equilibrium point at which both countries would be satisfied, at least with respect to privacy protection of European citizens when their personal information is transmitted over to computer and network infrastructures residing in US.

With respect to point 3, which is not unrelated to point 2, on 2 February 2016, the United States and the EU agreed on a new regulation framework for the transatlantic transfer of personal information of EU. This framework, named Privacy Shield, became part of the EU legal system following a favourable decision by the Commission on 12 July 2016. The EU-US and Swiss-US Privacy Shield Frameworks, was designed to create a collaboration framework to govern the dispatch of EU citizens' personal data to the US. It would enable organisations to handle data in the US and Europe in compliance with personal data protection requirements when transferring personal data from the European Union and Switzerland to the United States, for the support of transatlantic business activities. Two principal rationales underlay this agreement framework: (i) the economic rationale, for allowing legal transfer of personal information as part of transatlantic businesses, and (ii) the citizens' rights rationale, whereby EU citizens' personal information should not be collected, archived, and processed by US organisations in an inappropriate manner.

The Privacy Shield superseded the Safe Harbour US-EU agreement which was introduced as a result of a Commission decision on 26 July 2000. The transition from Safe Harbour to Privacy Shield was triggered by a CJEU mandate, dated 6 October 2015, which decreed that the Safe Harbour framework was inappropriate for regulating transatlantic transfers of EU citizens' data. Privacy Shield was a reaction by the EU administration to handle citizens' doubts and concerns of Europe's judicial system with respect to the sufficiency of Safe Harbour to protect EU citizens' personal data when transferred to non-European territories. However, the pan-European data regulators group, Article 29, criticised the Privacy Shield proposed by the European commission as a replacement of the Safe Harbour due to the absence of provisions for mass surveillance protection of EU citizens' data by US governmental agencies and authorities.

Moreover, Facebook recently failed in a last attempt to prevent a referral by Ireland's High Court of questions with respect to the legality of EU-US data transfer regulatory frameworks (Privacy Shield being the most prominent), to the CJEU. On 31 May 2019, Ireland's Supreme Court unanimously decided to reject Facebook's request to appeal with respect to the referral. This case has its origins in a dispute of Facebook with privacy lawyer Max Schrems on the use of another data transfer framework, called Standard Contractual Clauses (SCCs). Before this, he had contributed to questioning and invalidating the Safe Harbor agreement. In that case, he had succeeded in questioning the legality, after the 'NSA-Snowden Scandal' had made the headlines in 2013, of the personal data disclosures in US mass surveillance programs. The comeback to the Privacy Shield and SCC questioning is, thus, referred to as the 'Schrems II' case. It was the culmination of a series of serious doubts with respect to whether methods used by US companies for transferring and processing EU citizens' data really protect EU citizens' data from the US mass surveillance programs. On 9 July 2019, the hearing of the Schrems II case took place at the CJEU in Luxembourg. The principal entities involved in the case were the Irish Data Protection Commissioner ('Irish DPA'), Facebook Ireland Ltd. and the Austrian activist Max Schrems. The decision of Europe's supreme court is not expected before early 2020. However, both sides of the Atlantic are waiting anxiously since it could change radically the cybersecurity and privacy landscape in both US, Europe and their commercial relationship. It may even lead to radical transformations of the existing data transfer mechanisms, such as the Privacy Shield and the SCCs, impacting severely organisations and companies with limited means or alternative of securing transferred data.

Whereas the above is the more consolidated picture, stressing some notable negative incidents that emerged due to privacy protection regulation in Europe, the cited KPMG 2019 report envisages changes that are undergoing or will occur in the future both on data privacy (KPMG, 2018, pp. 8-9) and cybersecurity (KPMG, 2018, pp. 14-15). These changes will affect significantly, both, US and Europe as well as the transatlantic business collaboration perspectives. We summarise below this aspect.

The starting point is that there is unanimous agreement that the ongoing digital transformation and the high penetration of the Internet, globally, has brought as close as never before companies and consumers. There is, also, unanimous agreement that the privacy of organisations' and people's data should be protected at all costs. Disagreement begins, however, in the details and the implementation of privacy protection regulatory frameworks and technical measures.

As we discussed above, one of the main differences between US and European perspectives on privacy regulations is that in the US, due to the federated political system, there is no single regulatory framework across all US states. One consequence of this is that numerous companies exert influence on discussions according to their interests, in contrast with the EU where a single framework, the GDPR, was developed and unanimously accepted across all European countries.

Elements and principles of the GDPR philosophy have been adopted by several non-European countries, especially in Asia. There is also potential for worldwide adoption of GDPR under suitable adaptations which are already being considered in the US along with separate federal regulation frameworks. For instance, the DOC (US Department of Commerce) has issued a request for comments, on behalf of the central US Administration, with respect to a principles-oriented approach to consumers' privacy. The goal is to reach specific targets for the benefit of consumers such as data handling transparency, data control, user empowerment, data minimisation, data access and data correction by consumers (data owners), and accountability in case of a privacy breach. With respect to the enforcement of the regulation, it will be, in general, the jurisdiction of the FTC despite the fact that sectors such as banking and healthcare have their own regulations and enforcement agencies. The ultimate goal of the Administration is to develop a US-wide privacy regulation similar, in spirit, to the GDPR and the status it has in Europe.

In parallel, several discussions at the US Congress are targeting privacy and data protection issues, such as providing a federal privacy protection standard, possibly borrowing elements from EU's GDPR. This aims to reduce differences that arose out of the adoption of different standards in US states which has resulted in great differences in how privacy is handled, and breaches are managed across the US. Several players, among them some of the leading technology companies in the US, are pushing for such a federated privacy regulatory framework. For instance, as we mentioned before, California has already adopted AB 375, which is the California Consumer Privacy Act of 2018 (CCPA). This regulatory framework includes several elements of GDPR and is regarded as one of the strictest privacy regulations across the US. It will be applicable in California from 1 January 2020.

California's example is expected to be followed by other states. Many states have already enacted data breach notification legislation. But difficulties due to fragmentation remain in the US, while the EU's GDPR is applicable, on a global scale, for all organisations that handle personal data of EU citizens. Non-compliant organisations are liable to heavy fines. This, in turn, forces organisations to strengthen their privacy policies and be careful in which countries they operate. The ultimate goal is to raise privacy awareness among people so that they exercise their rights over the collection, use, and archiving of their personal information in line with regulatory requirements and law enforcement.

Although cybersecurity is considered as a top regulatory priority by regulators, there are no efforts for creating a uniform cybersecurity regulation framework across the US. Some efforts, nevertheless, have started working in this direction:

- **The US Administration has published a National Cyber Strategy while the Pentagon has, in parallel, published its own Cyber Strategy.** Several entities work within these directions including the Department of Justice, the Department of Homeland Security, and the Department of Defence. In addition, the National Institute of Standards and Technology (NIST), which belongs in the DOC, has established a Cyber Security Framework too.

- **Several federal financial service regulators have enacted cyber security requirements.** Some of them are, to some extent, in line with NIST's Cybersecurity Framework. Moreover, the federal banking agencies have released a notice of a proposed regulatory effort targeting improvements of cyber risk handling standards in 2016. However, there are no actions towards this direction yet.

- **Individual states have established cybersecurity regulations for their organisations.** For instance, New York's Department of Financial Services has released a set of Cybersecurity Requirements for companies active in the Financial Services sector. Elements of this effort include the appointment of a Chief Information Security Officer (CISO), the deployment of strong encryption standards for all non-public information, multifactor authentication (also based on biometric characteristics), breach notification policies and annual cybersecurity related reports.

4.2.3 ASIA

One of the most important initiatives in digital infrastructure regulation in Asia is the establishment of the Asia-Pacific Telecommunity, or APT, as the most widely accepted shaper among countries of the region. The APT was founded in 1979 through joint initiatives of the United Nations Economic and Social Commission for Asia and the Pacific (UNESCAP) and the International Telecommunication Union (ITU). The APT is an intergovernmental organisation that operates in conjunction with makers and users, including telecom service providers, manufacturers of ICT technology, as well as R&D insti-

tutions specialising in ICT and innovative digital technologies, as well as other organisations active in the field of communication, information and innovation. As of today, APT has 38 members from virtually all countries from the Asia-Pacific region. Most members are government agencies and ministries whose jurisdiction lies in ICT and its regulatory difficulties. APT also has 4 associate country members as well as 137 affiliate members (who are the makers but also users from the work of APT and its members) who are all strong industrial players active in the ICT among other sectors. For instance, for Japan, we see most of the ICT and digital infrastructure leaders as affiliate members.

The wide acceptance of APT in the region manifests its country-neutrality as well as effectiveness in handling digital infrastructure related issues. Under APT, we see the Asia-Pacific Forum on Telecommunication Policy and Regulation (PRF), which was established in 2001 and has now been recognised as one of the most important forums for regulators and policy makers active in the Asia-Pacific region. The PRF attracts around 100 high-ranking participants from the ministries and regulatory authorities of the Asia-Pacific countries each year. Over the years, the PRF has helped members strengthen their ICT policy and regulatory frameworks by sharing information, good practices and experiences, providing at the same time a forum for discussing key challenges related to regulating digital infrastructures and services.

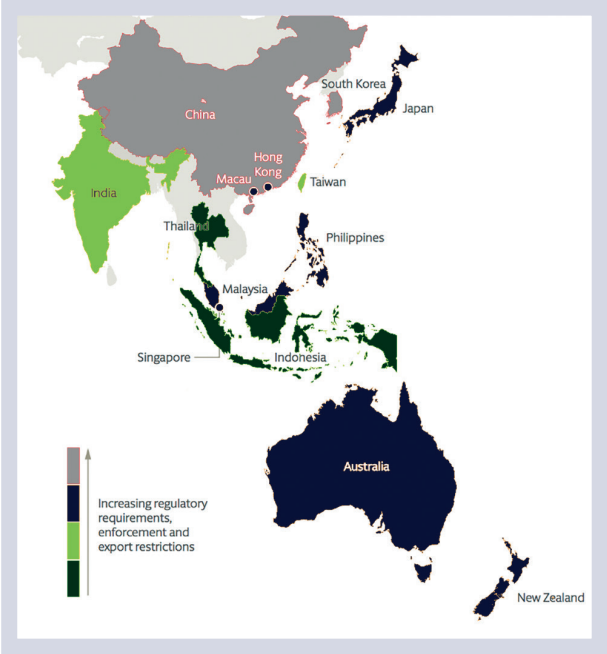


Figure 18: Regulatory landscape in Asia-Pacific, source: (Hogan Lovells, 2019)

With respect to specific countries, an important aspect of China's digital infrastructure regulatory framework is the exclusion of foreign technology. This may severely impact foreign security technology and security service providers wishing to expand their businesses, based on China-located infrastructure (e.g. networks), in China. The situation in Asia-Pacific is graphically summarised in Figure 19.

One of the most important cybersecurity regulatory advances in APAC in 2018 was India's strategic decision to introduce an all-embracing data protection legislation. The envisaged legislation will borrow many elements of the GDPR, such as a far outreach that affects foreign businesses that are active in India as well as the maximum fines, which may reach 4% of the data breach target's global turnover. China, also, introduced legislation based on GDPR as a reference anchor in order to dictate to organisations seeking business opportunities in China how to comply with the (often underspecified and unclear) data protection regulations in China's Cybersecurity Law as well as other data protection requirements. This is especially so with respect to data localisation requirements. In the privacy and security terminology, data localisation refers to the requirement of storing and archiving data on devices that are physically located within the borders of the country in which the data was created/collected. However, the legislation's data export review procedures have not yet been fully detailed.

Another important aspect of China's cybersecurity regulatory framework is the exclusion of foreign technology. This may severely impact foreign security technology and security service providers wishing to expand their businesses in China in which there are strong enforcement regulations with respect to storage of data. However, as of 2016 there is no all-inclusive regulation that describes when and under what conditions data transfers can (if at all) occur across borders. The situation is rather unclear since for certain data and industry types (e.g. handling governmental information) there are laws that strictly forbid cross-border transfer and storage while company data is not explicitly forbidden to transfer and store outside China.

As another important development in the APAC region, shortly after the introduction of the GDPR in the EU, in July 2018, an equivalency agreement was established between Japan and the EU. Under this agreement, Japan agrees to accept and apply EU data protection standards to personal data which is imported from EU countries.

South Korea's data protection and privacy protection regulations are considered among the most stringent in the APAC region as well as worldwide. The regulatory framework of the over-arching Personal Information Protection Act and the IT Network Act are supplemented by domain-specific legislation acts resulting in a very demanding compliance terrain for ICT players operating in this country. With respect to South Korea's regulatory framework, it is interesting to note that the requirement for data breach notification is meant to be a 'leakage' requirement, which implies that any unauthorized disclosure of personal data is a notifiable breach. On the contrary, a 'harm-based' requirement implies a no-

tifiable breach only if the disclosure can cause harm to the data owners/subjects.

More than any other country of the APAC region, the influence of GDPR is most markedly observed in India's draft Data Protection Bill, as detailed in a report titled A Free and Fair Digital Economy – Protecting Privacy, Empowering Indians. India's decision to introduce a data protection regulation is a significant advance in the APAC region as India is one of Asia's most populated nations with commensurate personal data residing in India's ICT infrastructures.

In conclusion, the impact of the GDPR for the APAC region is much further reaching than being a formal compliance requirement for Europe-APAC transfer of data and businesses. Legislation experts and data protection authorities of APAC countries are closely studying the GDPR with a view towards reshaping their own regulatory frameworks to encompass the most appropriate, to their needs, GDPR elements.

4.2.4 IDENTITY MANAGEMENT

Interactions on the Internet have raised concerns over centralised identity management systems, including:

1. Individuals rely on a central authority to assign and verify their identity online.

2. Individuals rely on a central authority to keep safe their personal information that is used to verify their identity.

3. A central authority can monitor the online transactions of an individual based on the identity verification requests that it receives for that individual.

Such concerns have led the EU to adopting the eIDAS regulation aimed to provide trust in cross-border exchange of ID data and to private proposals for self-sovereign identity (SSI) and standardisation by the Decentralised Identity Foundation on the Internet, where individuals can have complete control over their identity and over the information that can be used to verify it. The Sovrin project provides a blockchain-based platform to support SSI. In Sovrin, an individual can have different identifiers in relationship to different parties with which they interact, but companies have public identifiers; when two companies are involved in an online transaction with a user, the user can use separate identifiers for the part of the interaction with each company so that they cannot combine or share data about the individual (Morris, 2018).

At the same time, there is standardisation activity in the W3C on Decentralised IDs (DID). Here, DIDs are URLs that can be resolved to documents controlled by the owner of the DID; those documents provide cryptographic information, verification methods and service endpoints that can be used for verification. There is wider activity on standardisation and deployment of decentralised identity management systems under the auspices of the Decentralised Identity Foundation (DIF) which enjoys wide membership from the industry. They aim to support interoperability and foster the development of an ecosystem. They also provide a forum for the large community working on decentralisation, as evidenced by the events that they list on their website.

Issues of identification may also be mitigated by methods of de-identification. A "dataset is said to be de-identified if elements that might immediately identify a person or organisations have been removed or masked" (Bishop, p. 5). But this general picture of de-identification is made problematic by the fact that identifiability is being seen increasingly as a continuum rather than in binary terms. Especially as risks of identity disclosure are related to dimensionality of data (i.e. number of variables), how multiple data sources are linked, and to the computing power of data analytics architectures (ibid.). Consequently, while the risks of identity disclosure can be mitigated, these risks cannot be entirely eliminated by de-identification (ibid.). The risks of identification and whether or not users are comfortable in their use of the Internet can be illustrated in terms of the percentage of users who are aware of cookies tracking their Internet activity in Figure 20.

4.3 Platforms and Data Flow Imbalances

As seen, the growth of platforms has led to worries of data abuse, privacy violation and proper distribution of profit generated by data (Lee et al., 2017). An article published in 2014 in Time brought to the attention of the public various controversial ways in which Facebook uses the data of its users (Luckerson, 2014): tracking user movements or using user data in ads without consent. According to this article, Facebook paid more than $20 million for lawsuit settlement by disgruntled users. There is also reported danger of data abuse by platform users, partners or employees (reported in Lee at al., 2017). Data governance within platforms is complex for there are multiple parties contributing, deriving and using data complicating ownership, access, usage and profit-sharing of collected data and derived data (through data transformation/analysis).

According to Gawer (2009), certain types of platforms can function as the building blocks upon which an array of firms can develop complementary products, technologies or services to innovate. A distinction between intermediation-driven and innovation-driven platforms can be derived from the EIT Digital study on policy options for the platform economy (2019, pp. 7-8). In that report the following three types were identified:

- **Transaction platforms** facilitate exchange or transactions between different users, buyers, or suppliers. Typical exa-

mples are Uber, Airbnb, eBay, and also digital labour markets matching employers and workers (i.e. Upwork, Amazon Mechanical Turk, TaskRabbit).

▪ **Innovation platforms** facilitate players loosely organized into an innovative ecosystem to develop complementary technologies and products or services.

▪ **Integrated platforms** facilitate both transactions and the emergence of an innovation ecosystem. The typical example is Apple, which has both matching platforms like the App Store and a large third-party developer ecosystem that supports content creation on the platform. Other examples are Google, Facebook, Amazon, and Alibaba.

In view of the intermediation hierarchy proposed by Faravellon et al (2016), referenced in Chapter 1, we can conclude that some of the players included in the integrated type qualify as platforms intermediating abstract services. On the other hand, none of the dominant platforms qualify as truly open innovation-driven ecosystems. This latter type includes small and emerging ecosystems such as the Ocean Protocol Foundation new platform and SOLID. Large intermediation platforms, however, almost monopolise access to data, as we show below reporting from the above cited study by Faravellon et al (2016). This lead us to the implications of a data-driven economy for the long-term economic development of countries that is based on an essay by Weber (2017). Some examples are provided in the boxes below.

The US and China dominate the data-rich intermediation layer, whereas France and the UK show up mostly in the production layer (Figure 21).

The Figure 22 shows the power law of traffic dominated by Google and Facebook. The top 25 platforms attract most of the visits and, most likely, most of the data. They are thus major economic powers. For instance, in 2013 Amazon was larger than the next dozen Internet retailers combined.

Lastly, Figure 23 shows that US platforms receive traffic and data from most countries, whereas other countries struggle to limit traffic domestically.

Taking France, the graph above tells us that most of the platforms belonging to the French top 25 in terms of traffic are foreign. Only 22% of the national traffic is on national platforms. Overall, most traffic from other countries goes to US sites, about a third to national sites, and a tiny portion to sites of third countries.

Looking at these statistics, one may ask the question that is the focus of Weber's essay: "Put simply, do data flow imbalances make a difference in national economic trajectories? If a country exports more data than it imports (or the opposite) should anyone care? Does it matter what lies inside those exports and imports—for example, 'raw' unprocessed data as compared to sophisticated high-value-added data products" (2017, p. 338). During the period 1945-1982, when the Import Substitution Strategy dominated the theory of economic development, the answer would have been that it made a difference, since exporting raw materials and importing finished products was considered as the path to economic decline. In the period 1982-2002 of the Washington Consensus the spread of ICT and reduction in transportation pushed to unbundle supply chains, move pieces behind borders and organise the pieces. Since 2007-2008, the allure of above idea reduced (global flows of all kinds, except data, have decreased and not back to pre-crisis level).

In the context of the new data economy two perspectives would argue that the question above does not matter. First, that of the absolute gains from data flows claiming that what matters is being part of such flow. The McKinsey Global Institute (MGI) has put forward a very clear articulation of this position, arguing that directionality and content is irrelevant because data flows "*circulate ideas, research, technologies, talent, and best practices around the world*."(MGI, 2016). The second supported by Silicon Valley claims that 'open is best' which has been put forward vociferously whenever the EU has introduced or attempted to introduce regulation of platforms affecting US tech giants. (Kennedy, 2015) or introducing the digital services tax (ITIF, 2019, p. 3). A third position is what Weber calls 'data nationalism' as a sort of reflexive response and consists in trying to have their own data value-add companies 'at home' and to stop the new oil to flow abroad for the extraction of surplus (i.e., through data localisation laws or provisions within law).

Although Weber does not embrace data nationalism, through a series of thought experiments (see next two boxes below) he argues that a sort of new digital import substitution strategy is possibly the only alternative for a mid-sized country that is developed by data-peripheric and makes the example of France.

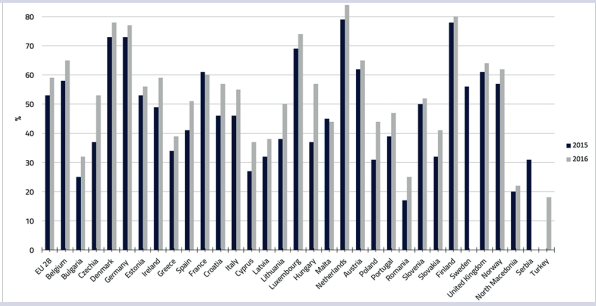


Figure 19: Users who are aware that cookies can track their internet activity, source: https://ec.europa.eu/eurostat/data/database

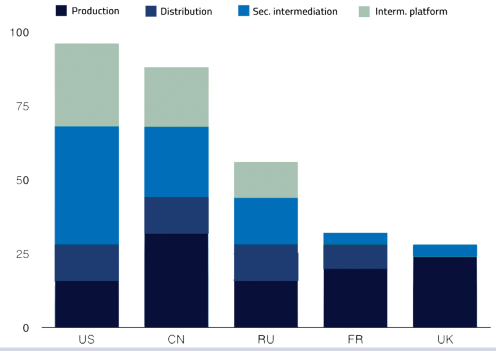


Figure 20: Proportion of national corporations at each intermediation level, source: (Faravelon et al., 2016, p. 27)

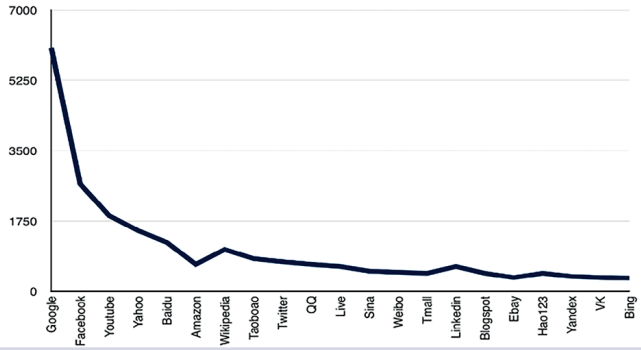


Figure 21: Proportion of national corporations at each intermediation level, source: (Faravelon et al., 2016, p. 27)

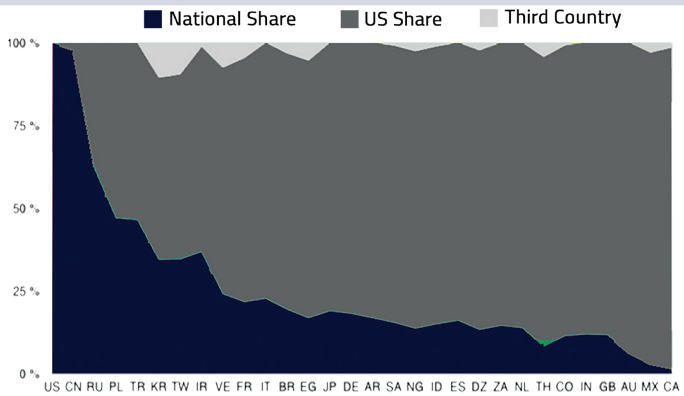


Figure 22: Top 25 platforms headquartered in the US, nationally, or in third country, source: (Faravelon et al., 2016, p. 29)

Country X passes a "data localisation" law requiring that data from X's citizens be stored in data centres on X's territory (ignore for the moment the various motivations that might lie behind this law). Now, a data-intensive transnational firm (say Company) has to build a data centre in Country X in order to do business there. The first-order economic effects are relatively easy to specify: Country X will probably benefit a bit from construction and maintenance jobs that are connected to the local data centre, while Company G will probably suffer a bit from the loss of economies of scale it would otherwise have been able to enjoy. It is the second and third order effects that need greater understanding. Imagine that the national statistics authority of Country X develops and publishes a "data current account balance" metric which shows that cross-border data flows two years later have declined in relative terms. Now the critical question: Is this a good or bad thing for X? Do nationally based companies that want to build value- add data products inside country X see benefits or harms? And does any of this matter to the longer-term trajectory of X's economic development?

Country X exports much of its "raw" data to the United States, where the data serves as input to the business models of intermediation platform businesses. Intermediation platform businesses domiciled in the United States use the "imported" data as inputs along with other data (domestic, and other imports from Countries Y and Z) to create value-added data products. These might be algorithms that tell farmers precisely when and where to plant a crop for top efficiency; business process re-engineering ideas; health care protocols; annotated maps; consumer predictive analytics; insights about how a government policy actually affects behaviour of firms or individuals (these are just the beginning of what is possible). These value-added data products are then exported from United States platform businesses back to Country X. Because the value-add in these data products is high, so are the prices (relative to the prices of raw data). Because there is no domestic competition in Country X that can create equivalent products, there's little competition. Because many of these data products are going to be deeply desired by customers in Country X, there's a ready constituency within Country X to lobby against "import restrictions" or "tariffs." And unless there's a compelling path by which Country X can kick-start and/or accelerate the development of its own domestic competitors to US platform businesses, there may seem little point to doing anything about this imbalance.

Each Uber ride in Paris produces a quanta of raw data—for example about traffic patterns, or about where people are going at what times of day—which Uber collects. This mass of raw data, over time and across geographies, is an input to and feeds the further development of Uber's algorithms. These in turn are more than just a support for a better Uber business model. Other, more ambitious data products will reveal highly valuable insights about transportation, commerce, life in the city, and potentially much more. If the Mayor of Paris in 2025 decides that she needs to launch a major re-configuration of public transit in the city to take account of changing travel patterns, who will have the data needed to make good decisions? The answer is Uber, and the price for data products that could immediately help determine the optimal Parisian public transit investments would be (justifiably) high.

Given costs of labour in the San Francisco Bay Area, there could be space for outsourcing of lower skilled tasks in the data value chain. Firm T in San Francisco contracts with Firm Y in France for doing such tasks. Is this a realistic rung on a climbable? Firm Y is at a huge disadvantage as it lacs access to all the data raw materials that would enable the jump. Firm T, in fact, is likely to distribute the outsourced work across multiple geographies. Suppose the French government try to push back by passing a law that requires more value-add data processing to take place France. Firm T in San Francisco would most likely respond by moving its data cleaning operations elsewhere, outside of country F. This is an attractive arbitrage play in data more so then ever, because investments in fixed capital for T's outsourcing operations are minimal to zero.

4.4 Cyber Security

The box below provides selective evidence on the global increase of cybersecurity breaches, their costs, and especially how they hinder the full development of the data economy. A disclaimer, however, is in order given the limited reliability of data on security breaches pointed out in a recent article (Florencio & Herley, 2011)[105].

Total breaches 2014-2016

- 2014: 1523 (with more than 10 million identities exposed: 11; total identifies exposed: 1.2B);

- 2015: 1211 (with more than 10 million identities exposed: 13; total identifies exposed: 564M);

- 2016: 1209 (with more than 10 million identities exposed: 15; total identifies exposed: 1.1B);

- In the last 8 years more than 7.1 billion identities have been exposed in data breaches.

Global estimates

- The likely annual cost to the global economy from cybercrime are estimated in more than $400 billion;

- Hundreds of millions of people having their personal information stolen cost as much as $160 billion per year;

- As cybercrime have impacts on export related jobs, Europe

could lose as many as 150,000 jobs due to cybercrime or about 0.6% of the total unemployed.

Costs to firms

▪ The 2015 Information Security Breaches Survey conducted in the United Kingdom showed that 90% of large organisations and 74% of small and medium-sized businesses reported they had suffered from an information security breach;

▪ For companies with more than 500 employees the average cost of the most severe breach was between €1.86 million and €4.01 million;

▪ For SMEs it oscillated between €95,840 and €397,1675.

Hindrances to Open and Big Data Economy

▪ The potential for data-driven innovation, provided cybersecurity is achieved, is a two-fold source of economic growth. First, directly as a new market with great economic potential of generating revenues by itself; Second, as a way of increasing efficiency and reducing administrative bottleneck;

▪ In the EU, if all framework conditions were in place, the EU data economy could increase up to EUR 643 billion by 2020 to EUR 272 billion in 2015.

Source: Total breaches 2014-2016 (Symantec, 2017); Global estimates (CSIS, 2014); Costs to firms(PwC, 2015); Hindrances to Open and Big Data Economy (OECD, 2013)

The rise of cybersecurity threats can be attributed to the expansion of the attack surface determined by nearly universal ICT usage and the take-off of the Internet of Things. In general, we observe an oxymoron with respect to cybersecurity. Everyone knows about rising rates of cybercrime even from the daily news. All companies handle sensitive information in their ICTs which can be a target of a cyberattack (which are on the rise, as a fact). Yet, cybercrime is still on the rise, which implies that not enough is being done. The explanation of this situation can be found in the behavioural bias and information asymmetry described in Section 2.1, and to fragmentation in the regulatory approaches both within and beyond the EU. First, firms may be not aware or will engage in liability dumping. Furthermore, they have to navigate complex multi-layered and fragmented regulation both in the US and in the EU (Aggarwal & Reddie 2018b; Timmers, 2018). According to Timmers (2018), the prioritisation of national security or national interests over common EU interests is a source of market failures. Fragmentation of the internal market means that EU countries impose different approaches (formal or informal) such as for ICT security requirements and standards[106]. In general, countries tend to enforce their own national direct cybersecurity regulations depending on their political, social, and financial agendas.

As illustrated in Deloitte's European Cyber Defense Report 2018 (Deloitte, 2019), all countries in Europe have already introduced at least one cyber security strategy on a national level targeting strong cybersecurity measures and high cyberattack resilience levels. To implement their strategy in cybersecurity, countries establish their own cybersecurity related organisations for developing regulations and responding to cyberattack incidents. This is exactly where a major critical issue in the global cybersecurity landscape emerges: a global cybersecurity viewpoint shared by all nations (much like the United Nations attempts to set common welfare and social standards across the globe) and a global cyberattack resilience strategy. A special issue of Journal of Cyber Policy (volume 3, issue 3, 2018), edited by Aggarwal & Reddie (2018a), provides an overview of cybersecurity state of play in several countries (US, China, Japan, Taiwan, France, UK, and Finland) and on the EU (2018), which is used below to highlight selectively some relevant aspects.

There are several specific security issues concerning **5G**, **IoT**, **critical infrastructures** and **AI**.

The coordinated risk assessment report carried out for Europe (NIS Cooperation Group, 2019; we refer below particularly to pp. 9-15) identifies the main security issues of **5G**. First, the value chain of 5G includes several stakeholders (mobile network operators, suppliers of mobile network operators, manufacturers of connected devices, service and content providers, and end-users) and this is one of the sources of security risks. Mobile network operators will play a key role, but many other players enter into the picture. Second, related to the previous source, new technical features (a move to software and virtualisation through 'Software Defined Networks (SDN) and Network Functions Virtualisation (NFV) technologies; 'Network slicing'; Mobile Edge Computing) bring new security challenges. They increase the complexity of the supply chain for they will force operators to rely heavily on integrators and other third-party suppliers, creating also a more articulated and potentially fragmented distribution of responsibilities. Third, functions currently performed physically and logically separated will move closer to the edge of the network. If not managed properly, these new features are expected to increase the overall attack surface and the number of potential entry points for attackers.

With **IoT** there is danger implicit in increasing interconnectedness of hardware and software utilised by businesses and governments in their adoption of IoT technologies and applications (Walport, 2014, p. 20). As the value of the IoT is the data extractable from its functions, any security weaknesses that are exploited can be economically and personally costly (i.e. in terms of loss of important assets or infringement of privacy). IoT has brought new risks. This applies in particular to consumer IoT, as it can involve 'non-technical' or 'uninterested' consumers, who connect an increasingly wide variety of devices to their home networks. They risk losing track of which devices are connected to the Internet over time, there-

fore making the efforts of securing them even more challenging. Connectable home devices, such as TVs, home thermostats or home alarms, create multiple connection points for hackers to gain entry into IoT ecosystems, access customer information, or even penetrate manufacturers' backend systems[108]. It also relates to the fact that various sectors and industries depend heavily on ICT components and the interdependence between current and future infrastructures (e.g. in smart cities environments, connected cars, energy smart grids).

As summarised by the European Parliament (European Parliament, 2019a), there is a specific need to focus security efforts on **critical infrastructures**. Energy and other utilities are increasingly controlled and monitored by networked industrial control systems. The electricity grids are being transformed into smart grids, in which more and more control functions are automated. This is expected to grow with the full deployment of 5G and IoT, which also potentially increases the risks. So, there is a clear risk that security breaches may paralyse critical infrastructure.

Hackers are becoming increasingly capable and are already probing and exploiting vulnerabilities in the energy system, as a number of incidents outside the EU have demonstrated[109]. The European programme for critical infrastructure protection (EPCIP), adopted by the Commission in 2006, established a framework for action aimed at improving the protection of critical infrastructure across all EU Member States and in all relevant economic sectors (European Commission, 2006). This was followed by 2008 Directive on European critical infrastructures (Council Directive 2008/114/EC) as the basis of the EU approach. In June 2019, the Commission published an evaluation of the Critical Infrastructure Directive (2008), which found that the Directive's relevance has diminished in the light of new and evolving challenges brought about by technological, economic, social, political and environmental developments (European Commission, 2019a). The Commission's recently adopted communication on cybersecurity in energy systems provides guidance. Binding rules for energy system operators are under development in the form of a new network code on cybersecurity (European Commission, 2019b) and a recommendation on cybersecurity in the energy sector (EU) 2019/553[110]. The recommendation underlines that the cybersecurity of the energy system, and notably the electricity grid, needs a dedicated sectoral approach because of real-time requirements, a mix of advanced and legacy technologies, and the cascading effects of disruptions. Experts see a growing need for improved exchange of knowledge and information, standardisation and certification, development of cybersecurity skills, and regulation. Further EU actions include:

- **Security of Gas Supply Regulation**: Regulation (EU) 2017/1938[111] deals with gas supply shortages caused by a number of risk factors, including cyber-attacks, war, terrorism and sabotage. It sets out rules for regional risk assessments and emergency planning, and introduces a mechanism for mutual assistance in the event of a severe gas supply crisis, based on the principle of solidarity;

- **Electricity Risk Preparedness Regulation**: Regulation (EU) 2019/941[112] is focused specifically on crisis prevention and crisis management in the electricity sector. It envisages the development of common methods to assess risks to the security of electricity supply, including risks of cyber-attacks; common rules for managing crisis situations and a common framework for better evaluation and monitoring of electricity supply security.

Finally, there are also cybersecurity implications for AI as it plays a role in risk management of cybersecurity (Timmers, 2019a). What are the ethical challenges in cybersecurity risk management, notably when making use of AI? Extensive monitoring and pervasive risk-prevention with the help of AI can be highly intrusive and coercive for people, whether employees or citizens. AI can also be so powerful that people feel that their sense of being in control is taken away. They may get a false sense of security too. Deep-learning AI is, as of today, not transparent in how it reaches a decision from so many data points, yet an operator may blindly trust that decision. AI also can incite freeriding as it is tempting to offload responsibility onto 'the system'. We are therefore confronted with a plethora of ethical issues when combining AI and cybersecurity in a risk management approach to strategic autonomy. They include erosion of individual autonomy, unfair allocation of liability, the fallacy of humans in the loop, the contestable ethics of mass surveillance and of trading off individual casualties versus collective protection.

As far as market structure goes, cybersecurity markets vary from monopolistic to competitive and fragmented structures. In China, the cybersecurity market is dominated by large monopolies with links to the national security apparatus (Cheung 2018). In Japan the role of the Ministry of Economy, Trade, and Ministry (METI), as well as a long-established practice of top-down policymaking have also contributed to the slow speed of growth in the cybersecurity sector (Bartlett, 2018). In the United States there is a plethora of companies marketing their cybersecurity programmes (Aggarwal and Reddie 2018b). In Europe markets are fragmented between a few large players and several small firms (Carr & Tanczer, 2018; D'Elia, 2018; Griffith, 2018; Timmers, 2018). Such structures are to a large extent shaped by the fact that national governments intervene on the basis of national security concerns. This is documented for the US (Aggarwal and Reddie 2018b), China (Cheung 2018), Finland (Griffith, 2018), France (D'Elia, 2018), and for EU as a whole (Timmers, 2018). In this respect, Timmers notes that in many European countries, cybersecurity suppliers developed through a close relationship to military and government buyers. The downside is a degree of national institutional dependency: "*Historically, industrial development in this area has been stimulated by governmental procurement and some highly innovative European companies*

in this sector are still largely dependent on this in their home country. A side effect of this situation is limited willingness for cross-border procurement, which is a barrier to the development of a common cybersecurity market" (European Commission, 2016b).

The EU cybersecurity policy has been developed in response to three drivers: preserving the internal market, combatting terrorism, and playing a global role (Timmers, 2018). It developed in several phases, the last of which started in 2013: the first fully-fledged EU Cybersecurity Strategy was launched, and a landmark EU cybersecurity law focused on economic resilience was proposed, the Network and Information Security Directive (NIS Directive)[113]. In 2016 an EU private-public partnership increased investment in research and innovation. Driven by the rapid rise of cyber incidents, this second phase can be characterised as moving towards a comprehensive and integrated EU cybersecurity policy.

Currently, we are in the third phase which started in September 2017 with an ambitious renewal of the overall strategy and several important legislative proposals. These include the EU Cybersecurity Act which introduces EU-wide IT security certification and an extended mandate for the cybersecurity agency ENISA, legislation for a common approach to scrutiny of foreign direct investment including for cybersecurity concerns, and legislation for strengthening EU cybersecurity competence. An EU meeting of all Heads of State also discussed cybersecurity. This third phase has been characterised by cybersecurity being Chefsache as a top political priority.

The most recent EU industrial policy argues that industry should become more adaptable, innovative and open to digitisation in order to be globally competitive (European Commission 2017a). EU cybersecurity industrial policy is thus firmly embedded in general EU industrial policy. In 2016 a further policy impetus to cyber-industrial capacity in Europe was given (European Commission, 2016c). As anticipated, the new EU cybersecurity policy foresees also scrutiny of FDI. The possible forms of industrial policy applicable to the cybersecurity domain are reviewed in Aggarwal & Reddie (2018a, pp. 6-8). The success or failures of industrial policies and of other types of policies depend on many factors in the design or implementation, including the possibly unforeseen strategic behaviours of the actors subjected to the policy.

An interesting case is that of breach notification laws aimed at market modification in the direction of incentivising firms to invest more in cybersecurity in order to avoid having to publicly report about the breaches. A quasi-experimental empirical study of the effects of California's law (introduced in 2002) found that while data breach notification laws have received considerable attention in recent years, their impact on firms' investment in web server security appears modest (Murciano-Goroff, 2018)[114]. But the Californian law did not include heavy fines. A theoretical principal/agent model shows that breach notification laws can produce social benefit (enough cybersecurity investments by firms to have positive overspill on economy and society) only if the fines foreseen are large enough (Laube & Bohme, 2016)[115]. This would suggest that the European GDPR and NIS Directive, both of which include sizeable fines, may be more effective laws.

NOTES AND REFERENCES

[1] It is the title of a piece in the Economist that extols the importance of data and concludes, however, with a critical note on competition (The Economist, 2017b). A Council of Europe note dated 18 November 2019 titled European Union as a hub for ethical data use states that the 'digital economy' has become a 'data economy', stressing the need to ensure trust in the current context of growing asymmetry (data concentrated in the hands of the largest players) and of steady concerns about privacy and security that "demand constant attention in order to maintain people's trust" (Council of Europe Doc 14070/19. Pp. 1-2). Earlier, on October 29, 2019, German Economy Minister Peter Altmaier presenting the European cloud project GAIA-X has been reported as saying that "Data are the resource of the future. That's why Germany and Europe needs data infrastructure that ensures data sovereignty and enables the sharing of data on a broader and secure basis" (https://www.straitstimes.com/world/europe/germany-to-unveil-european-cloud-gaia-x-to-rival-amazon-alibaba). That personal data as such is an unrefined raw material was certainly clear to the legislators drafting the GDPR (Regulation (EU) 2016/679) where article 4 introduce a wider definition of personal data clarifying that they include any information that can be used on its own or with other information to identify, contact, or locate an individual; Article 4 (1) recites that "Personal data means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".

[2] Through a large scale online discrete choice experiment the authors calculate the monetary value that users attribute to free online services. For instance, they report that the median Facebook user would ask about $48 to give up free access to this social network. So, the authors argue, in the digital economy many digital goods have zero price and as a result the welfare gains from these goods are not reflected in GDP or productivity statistics. Through this experiment they claim to have found that digital goods generate a large amount of consumer welfare that is currently not captured in GDP.

[3] This is summarised in a recent brief of the European Parliament on the economic impact of AI (European Parliament, 2019b), various studies project major productivity breakthrough thanks to the adoption of AI ranging from doubling annual global economic growth by 2035 (Accenture, reported in European Parliament, 2019, p. 3) to a 14% growth of global GDP by 2030 (PwC reported in European Parliament, 2019, p. 3) and to estimates that AI may deliver an additional economic output of around US$13 trillion by 2030, increasing global GDP by about 1.2 % annually (McKinsey Global Institute, reported in European Parliament, 2019, p. 3). AI is being utilised widely in a diverse range of domains, including monitoring traffic congestion, employee hiring, metering smart energy grids (Teich, 2019), and can produce important desirable results such as for instance analysing images to detect potentially cancerous cells (Al-shamasneh & Obaidellah, 2017), or helping predict where and when the next big earthquake will strike(Fuller & Metz, 2018).

[4] The UK Government Office for Science expects 100 billion connected IoT devices by 2020 generating 79.4 zettabytes (ZB) of data in 2025 and a global market value by 2020 estimated to be worth USD 14.4 trillion (Walport, 2014, p. 21).

[5] ETNO expects the number of mobile IoT connections in Western Europe is set to grow from 78.6 million in 2017 to 433.9 million by 2023.

[6] The large trove of data generated by IoT connections and devices will create fresh resources for growing data analytics and AI in Europe (Palovirta & Grassia, 2019). IoT-based health monitoring of patients with multiple chronic diseases in the Netherlands is reported has having generated 20 percent increase in efficiency (Rudas et al., 2019).

[7] From the German Economy Ministry website (see: https://www.bmwi.de/Redaktion/EN/Artikel/Digital-World/data-infrastructure.html).

[8] According to a market research report (Campbell et al., 2017), 5G will generate USD 12.3 trillion of global economic output by 2035 and that Investment in the value chain is expected to generate a further USD 3.5 trillion in output and provide support for 22 million jobs.

[9] https://blogs.microsoft.com/on-the-issues/2019/05/20/gdprs-first-anniversary-a-year-of-progress-in-privacy-protection/

[10] The Impact Assessment refers to the latest example of a ransomware cyber-attack in May 2017 shows the potentially massive impact of a cyber- attack across sectors and coun-

tries: more than 150 countries and over 230,000 systems were affected, including those related to essential services such as hospitals, despite the damage being contained this time in comparison to the potential (deeper) consequences it may have had (WannaCry Ransomware Outburst, Infonotes, ENISA, 2017: https://www.enisa.europa.eu/publications/info-notes/wannacry-ransomware-outburst). This example is just the last of a series: more than 4,000 ransomware attacks have occurred every day since the beginning of 2016, a 300% increase over 2015 (How to protect your networks from ransomware, CCIPS, 2016: https://www.justice.gov/criminal-ccips/file/872771/download) . Cyber incidents cause major economic damage to European businesses, undermine the trust of citizens and enterprises in the digital society and affect citizens' fundamental rights. A 2014 study estimated that the economic impact of cybercrime in the Union amounted to 0.41% of EU GDP (i.e. around EUR 55 billion) in 2013(CSIS, 2014); with Germany being the most affected Member States (1.6 % of GDP). A recent report, in the aftermath of the 'wannacry' attack, estimated that a serious cyber-attack could cost the global economy more than $120bn (£92bn) – as much as catastrophic natural disasters such as Hurricanes Katrina and Sandy (Counting the cost – Cyber exposure decoded, Lloyd's and Cyence, 2017: https://www.lloyds.com/~/media/files/news-and-insight/risk-insight/2017/cyence/emerging-risk-report-2017---counting-the-cost.pdf).

[11] This report, part of the Transatlantic Dialogue on Security and Freedom in the Digital Age and partly funded also by the Commission, identified a large number of technological sovereignty proposals put forward in Europe (see annex III) and produced an impact assessment against the OECD Principles for Internet Policy Making (http://www.oecd.org/internet/innovation/48289796.pdf).

[12] Timmers defines strategic autonomy as "the ability, in terms of capacity and capabilities, to decide and act upon essential aspects of one's longer-term future in the economy, society and their institutions" (Timmers, 2019b, p. 2). According to Timmers, this is the result of the interplay between international tensions (in relations with Russia and China and also transatlantic between Europe and the US), the growing dependency on digital technologies throughout economy, society and democracy and the rise of cyber-threats The in-house policy advisory body of the European Commission stated recently that digital technologies affect all elements of strategic autonomy (EPSC, 2019, p. 3). After tense NATO and G7 Summits in May 2018, Merkel said, "We Europeans must really take our fate into our own hands" (See https://www.reuters.com/article/us-germany-politics-merkel/after-summits-with-trump-merkel-says-europe-must-take-fate-into-own-hands-idUSKBN18O0JK, retrieved 15/10/2019, cited in Timmers, 2019a, p. 1). When European Commission President Jean- Claude Juncker gave his 2018 State of the Union speech last September with the title, "The Hour of European Sovereignty", he argued that the time had come for the EU "to become more autonomous and live up to our global responsibilities" (European Commission, The Hour of European Sovereignty, retrieved from: https://ec.europa.eu/commission/sites/beta-political/files/soteu2018-speech_en.pdf ,15/10/2019, cited in Timmers, 2019a, p. 1). In December 2018 18 EU countries jointly stated that the EU "must adapt its trade policy to defend its strategic autonomy", specifically referring to a range of fields including cybersecurity and AI. hey also said that the EU must "ensure its technological autonomy by supporting the development of a digital offer and create global reference players" (Friends of Industry 18 December 2018, retrieved from: https://www.bmwi.de/Redaktion/DE/Downloads/F/friends-of-industry-6th-ministerial-meeting-declaration.pdf?__blob=publicationFile&v=6 , 15/10/2019, cited in Timmers, 2019a, p. 1).

[13] ETNO cites that out of "Out of the Forbes Digital 100 ranking of the top 100 public companies shaping the worldwide digital economy, only 13 are from the EU28" (Palovirta & Grassia, 2019).

[14] As reported by Cohen (2016, p. 381-382), in U.S. legal and policy circles anti-European posturing figures high and takes vitriolic forms. As put by Cohen, the historical U.S. antipathy to European-style bureaucracy is not sufficient to account of the violence of anti-European statements (see various harshly anti-European media coverage of technology related regulatory action by the EU cited by Cohen, 2016, footnotes 30 through 32 at pp. 381-382). European regulators are charged of attempting to institute a regime of economic protectionism. Posturing putting U.S. first has also been enacted as legislation like the Cloud Act introduced in 2018 by the Trump administration requiring American firms to provide law enforcement with customers' personal data on request, even when the servers containing the information are abroad, which is seen as triggering the push for a European cloud by Germany (Stupp, 2019).

[15] https://edition.cnn.com/2019/11/01/tech/russia-internet-law/index.html ; see also https://data-economy.com/internet-iron-curtain-comes-down-across-mother-russia/ . Concern arising from this news is largely reflective of fears that this law and its enactment would make it easier for censorship and surveillance of politically sensitive information and views. This greater censorship would take place once tools such as Deep Packet Inspection (DPI) 'which involves data processing that looks in detail at the contents of the data being sent' similar to how the Great Firewall operates in China.

[16] See Most Valuable Companies in the World – 2019(retrieved from: https://fxssi.com/top-10-most-valuable-companies-in-the-world, 16/10/2019).

[17] The GAFAM are the five most valuable firms in the world and they collectively racked up over $25bn in net profit in the

first quarter of 2017 (The Economist, 2017b). Furthermore, oligopolistic settings may stifle competition and innovation and reduce benefits to society.

[18] Digital innovations still do not reach everyone, as digital divides still exist, especially as those "who lack safe and affordable access to digital technologies are overwhelmingly from groups who are already marginalised: women, elderly people and those with disabilities; indigenous groups; and those who live in poor, remote or rural areas" (United Nations, 2019, p. 11). Developments in digital technologies do not therefore take place in a vacuum, and the values that guide technological development must be set out clearly. The "application of technology must be aligned with investments in human capital, infrastructure and environmental protection" (United Nations, 2019, p. 15). Furthermore, besides the economic impacts of digital technologies, the social impacts also prompt regulatory attention. It is necessary to 'ensure that advances in technology are not used to erode human rights or avoid accountability' (United Nations, 2019, p. 14). With the increasing scale that digital technologies and digital services are reaching, collaboration between businesses and governments will be necessary as both have a duty to protect rights, provide solutions, evaluate risks and assess the impact of their actions (especially businesses) on human rights (Ibid.).

[19] This is underscored, for instance, in the earlier cited UK government report on IoT strongly recommending government, industry, and international partners to agree best practice security and privacy principles based on "security by default" to ensure trust (Walport 2014, p. 10).

[20] Trust and social capital and are two sides of the same coin as they both pertain to relations and the expectation entailed in them. The concept of social capital commands an ever-expanding body of literature that cannot be discussed here. General approaches define social capital in slightly different ways depending on the theoretical perspective (Bourdieu, 1986; Coleman, 1988; Coleman, 1990; Putnam, 1993, 2000). At a very basic level it can be said that the concept entails both a norms/values and instrumental dimensions within the domain of social networks. At macro level social capital can be equated to civic sense entailing norms, social values, trust, and social network (especially participation in association). It is, however, most important to concept of trust and in relation to social capital, for trust is fundamental for adoption and beneficial use of new technological possibilities. Trust is the social glue that enables collaborative and productive practices in the digital ecosystem. Scholars of trust distinguish between generalized and particularize trust (Couch & Jones, 1997; Delhey et al., 2011; Freitag & Traunmüller, 2009; Putnam, 1993, 2000; Stolle, 2002; Yamagishi et al., 1998; Yamagishi & Yamagishi, 1994). Particularized trust, also referred to as 'thick trust' (Putnam 2000) concerns a close network of social proximity (i.e. family and friends). Generalised trust is a more abstract attitude toward other people and expectations about their behaviours. It entails some implicit consideration of risk and uncertainty leading to a 'estimate' of the trustworthiness of others (Coleman 1990). In other words, generalised trust can be defined as an attitude entailing reliance on the benevolence of human nature (Couch & Jones, 1997; Yamagishi & Yamagishi 1994) or the attitude to give most people the benefit of the doubt (Putnam, 2000, p. 133). Generalized trust is, thus, a critical element of social capital and the foundation of civic behaviour (Stolle 2002), the basis of reciprocity and social connectedness (Delhey et al. 2011), and as a 'bridging' mechanism linking people to engage with others unlike themselves (Stolle & Hooghe, 2004). Obviously, since in the digital landscape transactions and interactions among strangers are crucial, generalised trust as a willingness to rely on 'abstract others' is crucial. Online exchange settings are characterised intrinsically by two forms of information asymmetry: the identity of online parties (Ba, 2001; Pavlou & Gefen, 2004)– which has three dimensions unidentified, anonymous, and not possible to bind to a single person – and the quality of the exchanged object (Gefen et al., 2008; Jøsang et al., 2007). The second source stems from the fact that the online consumer has no opportunity to see and test out the products before he/she purchases while payment usually occurs in advance. This knowledge gap between buyer and seller necessitates a high level of trust in the online context as compare to the analogic one. In the case of some very limited and particular instances, such as transactional and sharing platforms, one source of trust is the importance of reputation. Reputation reduces the asymmetry concerning the quality of the exchanged object by increasing confidence in the person offering it (Jøsang et al. 2007). This should be backed also by the utilitarian consideration that reputation is a 'value' that can influence the capacity to exchange or sell a particular good or service (Burnham, 2011). In online exchanges reputation is network produced by members referrals and ratings (Jøsang et al. 2007). When the parties are total strangers to one another reputation systems are collaborative filtering mechanisms helping the emergence of generalised trust (Corritore et al., 2003). In a way within a given community of online exchangers reputational ratings are a sort of social control by which the members police themselves (Abdul-Rahman & Hailes, 2000; Ba, 2001). Opportunistic behaviour is in principle sanctioned should not be imitated at least according to economic theories of cooperation(Axelrod, 1984). A second mechanism with the capability to increase trust in online marketplaces is the implementation of social networking features, or the leveraging of pre-existing relationships. Such integration two purposes in building online trust: confirming identity and establishing transitive trust (Hogg & Adamic, 2004; Jøsang et al., 2007; Kwan & Ramachandran, 2009; Swamynathan et al., 2008).

[21] On the peculiarity of the EU approach see also (EPSC, 2018).

[22] See: https://ec.europa.eu/digital-single-market/en/free-flow-non-personal-data.

[23] See: https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy.

[24] See: https://ec.europa.eu/commission/presscorner/detail/en/MEMO_19_2750.

[25] See: http://www.crenger.com/ovw2.html .

[26] See: https://simplicable.com/new/digital-infrastructure . Note that this and the one above are just two examples chosen selectively among dozens of definitions available on the web.

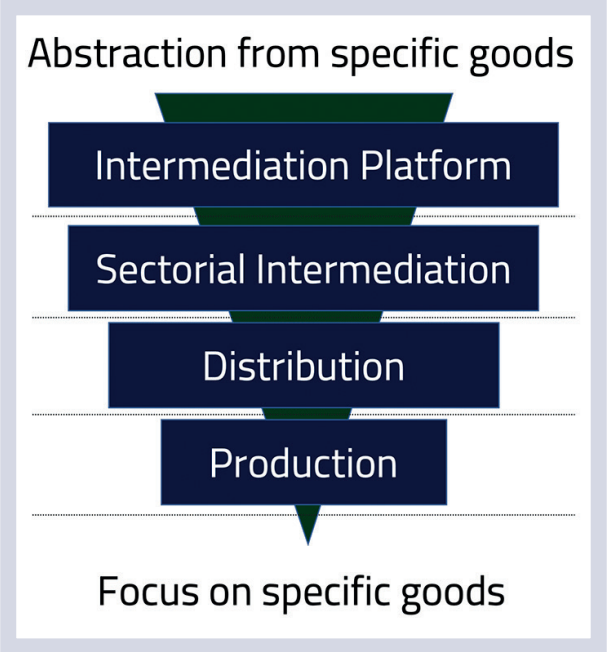
[27] According to Atkinson et al. (2016), for instance, infrastructure refers to system for the transportation of goods, people, and information and does not include the nodes (i.e. airports, oil terminals, electric generating stations) and, thus, cloud computing is not an infrastructure but a broad system of nodes This definition contrasts with a number of sources that consider cloud as a key digital infrastructure. A Global X piece, for instance, considers cloud computing as 'the' digital infrastructure that power the businesses of the future(Jacob, 2019). The earlier cited German project document on Gaia-X consider the cloud as key and strategic infrastructure where data sovereignty must be achieved, and dependency reduced (BMWi, 2019, p. 5). Atkinson et al digital infrastructure is defined as Information Technology systems that electronically collect, process, and transmit information. They can be of two kind: hybrid and dedicated. The former are traditional infrastructures with an IT component (i.e. smart grid), while the latter are just digital and self-standing (i.e. fibre optic cable to transfer digital Internet packets). Hybrid digital infrastructure are increasingly ubiquitous as a result of smart sensors and IoT and is a matter connected to the issue of critical infrastructure. A report by Arthur D. Little considers digital infrastructure as the key driver of competitiveness of the future a define it broadly to include from physical network to IoT and various other elements (Rudas et al., 2019). On the other hand, the UK government define digital infrastructure as networks enabling other infrastructures in driving economic growth and productivity and defined as targets full fibre coverage by 2033 and for the majority of the population to be covered by a 5G signal by 2027 (DCMS, 2018, p. 37). This review of diverging definitions could go on much further, but it is beyond our scope to enter into definitory disputes and technicalities. The World Economic Forum report on delivering digital infrastructure also takes a broader approach to the definitions including into it, not only clouds, but also IoT and social media platforms (World Economic Forum, 2014)

[28] For a general review of definitions, theory and empirics concerning infrastructures see Fourie (2006); for systematic review of the relevant scientific literature focussing on the impact of infrastructure on economic growth and competitiveness see Palei (2015), while for an example of concrete estimates see a report delivered for the US Congressional Research Services (Stupak, 2018);

[29] The earlier mentioned economic paper by DG ECFIN reports that ICT is regarded to account for 5% of GDP growth and 20% of productivity growth in Europe (Lorenzani & Varga, 2014, p. 7). The authors, using a modelling simulation, project that digital structural reforms that reinforce Europe infrastructure could have a 1% yearly impact on long term growth and deepened efforts could reach an impact of an additional 2.1% of GDP growth over the baseline (Lorenzani & Varga, 2014, p. 1). According to government estimates in 2017 digital infrastructure contributed £33bn to the UK economy (that is 1.8% of Gross Value Added), and increase of about 33% compared to 2010 (DCMS, 2018, p. 37). Increased connectivity and download speed in certain postcode areas created on average added £9bn turnover for firms in those areas (ibid.).

[30] As stressed by the coordinated risk assessment report carried out for Europe (NIS Cooperation Group, 2019, p. 9).

[31] One can make a distinction with respect to the degree of intermediation entailed in their business model. Following Faravellon et al (2016) one can envisage a hierarchy of intermediation as depicted in the figure below.



Source: Faravelon et al. (2016, p. 25)

At the lowest level we find production actors (i.e. media corporations producing contents) followed by distributors (i.e. Amazon or Netflix) and sectorial intermediation (i.e., LinkedIn). At the highest level, true intermediation platforms offer 'abstract services', such a social network offering various functionalities without focus on specific uses and allowing to build on top using its API. So, these platforms offer an ecosystem upon which other can build or distribute their services. Google and Facebook belong to this category, Apple and Microsoft do not, whereas Amazon may overlap several layers.

[32] Consider, for instance, the example of the data ecosystem for the self-driving car. According to estimation by RAND (Kalra & Paddock, 2016), 500 billion to 1 trillion miles driven are needed to get AI models accurate enough for production deployment of self-driving cars; it would be too expensive for a single car maker to generate that much data alone. Pooling data into a platform would make this operation more efficient and, if not monopolised by one single player, create opportunity for innovators. So, equalisation of access to data enables a larger pool of AI innovators to improve their systems, bring them to market and create value.

[33] Aggarwal & Reddie (2018a, p. 293).

[34] A broad definition of the costs of cybercrime include: a) criminal revenues, that is the gross receipts from crime (this is a cost to society), the other three are costs for victims and also to society; b) Direct costs: value of losses, damages experienced by victim (but also costs of ex post fixing the problem: the costs of cleaning systems from malware; c) Indirect costs: values of losses and opportunity costs imposed on society (i.e., loss of trust, foregone sales, reputational costs for individual firms or entire industries, etc.); d) Defence costs: values of prevention efforts (Anderson et al., 2019). Going more granularly as many as 14 different type of costs (Deloitte, 2016). It is important to stress that indirect intangible reputational costs may be even greater than direct costs, when measured by the impact on stock markets for listed companies (Cavusoglu et al., 2004). In this respect, it has been recently shown that a broad based approach is needed to quantify such costs (Haislip et al., 2019). These authors explain that one should not stop observing that at times there is difference between media attention/regulatory activism on one hand, and lack of response by capital markets on the other. They argue that a broad-based analysis of cybersecurity breaches costs, not focusing only on targeted firms but examining the effect on non-breached industry peers through the lens of capital markets, auditors, and affected insurers Non-breached peers experience significant negative equity returns around the announcement of a cybersecurity breach in their industry, together with a material increase in audit fees during the year of the infraction. Also, significantly negative equity returns for insurers with material cybersecurity exposure can be observed.

[35] In the well-known case brought by the Spanish Data Protection Authority against Google Spain (see in depth analysis in Zuboff, 2019, chap. 2, Section V) the Court of Justice of the European Union ruled against Google Spain in 2014 in a case brought by a man who wanted outdated information about him removed from Google's search results (Google Spain SL and Google Inc v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González, 13 May 2014; Doc C-131/12, ECLI:EU:C:2014:317 CJEU). The ruling limited the extension of the rights to be forgotten within EU jurisdictions.

[36] Citing Bamberger & Mulligan (2015), O'Hara & Hall (2018) suggest that the law may promote a box-ticking mentality that may turn out to be no more effective than the tort-based approach of the United States.

[37] See: https://ec.europa.eu/commission/priorities/stronger-global-actor_en

[38] ITIF argued against any regulation of platforms when this was envisaged in the EU (Kennedy, 2015), placed Europe as a whole in its 2018 black list of the Ten Worst Digital Protectionism and Innovation Mercantilist Policies for having "Attempted to introduce a mercantilist digital services and digital profits tax that would have targeted U.S. tech firms almost exclusively" (ITIF, 2019, p. 3), and Germany in same black list of 2016 introducing "forced local data-storage requirements—ostensibly due to privacy and cybersecurity concerns—as part of a new telecommunications data law"(ITIF, 2017, p. 2).

[39] By 2025, the GSM Association (GSMA) expects 5G connections to reach 1.1 billion, some 12 per cent of total mobile connections. It also forecasts overall operator revenues to grow at a CAGR of 2.5 per cent, to reach USD1.3 trillion by 2025 (Reported in ITU, 2018, p. 4). One report estimates that 5G will generate USD 12.3 trillion of global economic output by 2035, with the greatest growth in sales activity coming from manufacturing because of an anticipated increase in spending on 5G equipment. This is followed by sales growth in the ICT sector driven by higher expenditure on communications services. Investment in the value chain is expected to generate a further USD 3.5 trillion in output and provide support for 22 million jobs by 2035 (Campbell et al., 2017). Another market research company estimates that that mobile broadband operators will reap 5G revenues of $247 billion in 2025 with North America, Asia-Pacific, and Western Europe being the top markets. In 2014, the European Commission had estimated that the total cost of 5G deployment across the 28 Member States will be EUR 56 billion, resulting in benefits of EUR 113.1 billion per annum arising from the introduction of 5G capabilities, and creating 2.3 million jobs. It is also estimated that benefits are largely driven by productivity in the automotive sector and in the workplace generally. Most of the benefits are expected in urban areas while only 8 per cent of benefits (EUR 10 billion per annum) will be realized in rural areas(DG CONNECT, 2014).

[40] See: ABI Research projection at: https://www.abiresearch.com/press/abi-research-projects-5g-worldwide-service-revenue.

[41] These concerns are supported, as the estimated cost to deploy a small cell-ready 5G network – assuming fibre backhaul is commercially feasible – can range from USD6.8 million for a small city to USD55.5 million for a large, dense city (2019, p. 68). Given the considerable CAPEX investment required in deploying 5G, operators face major challenges in making the investment case for 5G.

[42] Barcelona's smart street lights, which analyses required brightness via the IoT and have contributed to a 30 percent energy reduction; a networked and intelligent transport system on the M42 motorway in the UK, which has decreased travel time by 25 percent and accident frequency by 50 percent; and IoT-based health monitoring of patients with multiple chronic diseases in the Netherlands, which has led to a 20 percent increase in efficiency (Rudas et al., 2019).

[43] Underpinned by 5G, the number of mobile IoT connections in Western Europe is set to grow from 78.6 million in 2017 to 433.9 million by 2023 (Palovirta & Grassia, 2019); the growth in connected IoT devices is expected to generate 79.4ZB of Data in 2025, according to a IDC Forecast 18 June 2019); Three quite different projections of IoT market value by 2020 (reported in Walport 2014, p. 19 p. 21):

- Cisco: US $ 14.4 trillions. Cisco further identifies increases in employee activity, reduction of costs, improved citizen experience and increased revenue as four drivers in the public sector following implementation of IoT, estimating 'over 25% of an estimated $19 trillion global market value [would be] available up to 2022' in the public sector (p. 21). Consequently, a 'future enriched by the Internet of Things is likely to be one where good security practice supported by robust system design is an essential part of everyday life';

- IDC: US $ 7.1 trillions;

- Gartner: US $ 1.9 trillions;

[44] Atali et al (2019, p. 2) point out in a McKinsey article that by 2021 "about 35 percent of all enterprise workloads will be on the public cloud, and 40 percent of companies will use two or more infrastructure-as-a-service(IaaS) and software-as-a-service (SaaS) providers".

[45] Cloud computing delivers IT services directly over the internet without any concern for the interoperability on-premises. Traditional on-site computing architectures, on the other hand, are "capital- and time- intensive without a high degree of scalability" (Jacobs, 2019, p. 2). Businesses that utilise public cloud systems due to the technical advantages, will also benefit from economic advantages due to the cost efficiencies from economies of scale (Jacobs, 2019, p. 4).

[46] An instance of this is in Slovenia. While for a long time "public administrations databases are unlinked and located in different silos", these administrations are attempting to work towards developing "a common governmental platform for data analytics, which will integrate a data warehouse, a data lake for big data and business intelligence and artificial intelligence functionalities"(Battisti et al., 2019, p. 6).

[47] See: https://solid.inrupt.com

[48] A debate is focussing one legally requiring AI to be transparent. Such requirement needs a strict and clear definition of AI. The regulation of technology is useful if there is something essential about it that is not regulated by already existing laws and regulations. Calo specifies this argument in the context of AI, proposing to identify the three essential qualities of AI (2015): embodiment (interaction with physical world), emergence (unpredictability of interaction with its environment) and social valence (whether people treat AI agents as human beings). Surely AI meets these criteria and, thus, would require new laws and regulation, but seemingly the problems of regulating AI is that there is no accepted definition of what AI is (Scherer, 2016). As illustrated by (Buiten, 2019, pp. 3-5), most definitions are circular as for instance those focussing on intelligence and autonomy, where then there is not a clear-cut definition of these qualifying terms. So, Buiten advocates looking beyond the opaque concept of AI, focusing on the concrete risks and biases of its underlying technology: machine-learning algorithms and consider what it means requiring them to be transparent and the right to explanation (2019).

[49] For instance, AI applications are analysing images to detect potentially cancerous cells(Al-shamasneh & Obaidellah, 2017). They can help predict where and when the next big earthquake will strike(Fuller & Metz, 2018). On the other hand, Microsoft's chatting bot Tay had to be shut down after 16 hours because it became racist, sexist, and denied the Holocaust (See: https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism/), errors in credit checks, recidivism (Teich & Tirias Research, 2018), and various other examples that are fuelling a variety of concerns about the accountability, fairness, bias, autonomy, and due process of AI systems (Pasquale, 2015; Ziewitz, 2015).

[50] Bias can arise in algorithms in several ways. First, the data we have collected may have been preferentially sampled, and therefore the data sample itself is biased.(Olhede & Wolfe, 2018)Second, bias can arise because the collected data reflects existing societal bias.(Caliskan et al., 2017) To the extent that society contains inequality, exclusion or other traces of discrimination, so too will the data.(Goodman & Flaxman, 2017)For instance, differences in arrest rates across racial groups may be replicated by algorithm calculating recidivism risk.(Chouldechova, 2017) Another example could be underrepresentation of women in particular jobs, from which a hiring algorithm may derive the rule that men are preferable candidates (See for instance 'Amazon scraps secret AI recruiting tool that showed bias against women' (Reuters, 10 October 2018: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-thatshowed-bias-against-women-idUSKCN1MK08G). In short, machine learning can reify existing patterns of discrimination (Goodman & Flaxman, 2017, p. 3). Output may also be biased if the training data is not representative of the real-world environment in which the system is supposed to perform. An algorithm may

produce unexpected and undesirable results if it encounters a situation that is markedly different from the data it was trained on. As algorithms become more sophisticated, their decision-making process may become less tractable. The problems associated with biases in data and insufficient testing may get more pervasive as the decision model gets more complex, because these biases may be more difficult to predict or identify. For instance, avoiding biased results rooted in social inequalities is difficult if sensitive information, such as ethnicity, is correlated with seemingly neutral variables, such as home address. In such cases, removing the sensitive variable will not prevent the biased result. Sophisticated algorithms may be able to reconstruct sensitive information from other inputs, even if they are not given this information (Doshi-Velez & Kortz, 2017). With sufficiently large data sets, the task of exhaustively identifying and excluding data features correlated with 'sensitive categories' a priori may be impossible. If we are not aware of correlations between variables, these hidden relationships may obscure the rationale for how predictions are being made (Olhede & Wolfe, 2018, p. 4; Goodman & Flaxman, 2017, p. 4). The chosen decision model may also turn out to be unsuitable if the real-world environment behaves differently from what was expected. If an algorithm is fed with input from users or consumers, designers may have to account for the interaction with the social-learning abilities in the decision model. Microsoft learned this lesson the hard way in 2017, when its chatting bot Tay had to be shut down after 16 hours because it became racist, denied the Holocaust and supported Hitler. Tay was intended to be empathetic and was highly successful at that. However, this meant that as Twitter users started deluging Tay with racist, homophobic and otherwise offensive comments, Tay stopped being family-friendly and its type of language changed dramatically. Since then, Microsoft launched a new bot that was programmed differently with respect to its responses to input.

[51] AI Global Surveillance Index (AIGS) is an initiative of the Carnegie Endowment for International Peace and all reference material can be found at: https://carnegieendowment.org/files/AIGlobalSurveillanceIndex.pdf ; An interactive map keyed to the index that visually depicts the global spread of AI surveillance technology can be accessed at: https://carnegieendowment.org/AIGlobalSurveillance . There is also an open Zotero library with all reference source material used to build the index. The methodology for the construction of the index is illustrated in Feldstein (2019, pp. 5-7 and pp. 25-28). Here it suffices to say that the index, based on the compilation of a vast array of sources, includes detailed information for seventy-five countries where research indicates governments are deploying AI surveillance technology and breaks down AI surveillance tools into the following subcategories: 1) smart city/safe city, 2) facial recognition systems, and 3) smart policing.

52 See: European Parliament Committee on Legal Affairs, Civil law rules on robotics (2015/2103 (INL)), p. 10.

[53] Guidelines for achieving this framework have been addressed to all relevant stakeholders including 'companies, researchers, public services, government agencies, institutions, civil society organisations, individuals, workers, and consumers. These guidelines have the following seven key requirements, which these stakeholders can voluntarily follow: human agency and oversight (to mitigate any infringement of fundamental rights and improve explainability), robustness and safety, privacy and data governance (to ensure citizens have control of their data), transparency, diversity, non-discrimination and fairness, societal and environmental well-being, and accountability. These requirements ensure that human values are applied to the development of algorithms that increasingly play a role in our daily lives, especially as they force us 'to rethink our understandings of human dignity and agency' as they 'are increasingly sophisticated at manipulating our choices - for example, to keep our attention glued to a screen'. (United Nations, 2019, p. 24). Greater oversight will make it more possible to uncover instances of discrimination within algorithms before profound real-life consequences (United Nations, 2019, p. 25). This is why there are a growing number of initiatives, such as the Institute of Electrical and Electronics Engineers (IEEE)'s Global Initiative on Ethics of Autonomous and Intelligent Systems. Such initiatives assess the delegation of human responsibility and legal accountability in the design of autonomous intelligent systems. These initiatives are important because these systems 'raise the danger that humans could evade responsibility for decisions made or actions taken by technology they designed, trained, adapted or deployed' (United Nations, 2019, p. 25).

[54] This data perspective is based on a number of data ethics principles, such as: foresighted responsibility (concerning network effects, effects of scale and changing actor constellations), respect for the rights of the parties involved, data use and data sharing for the public good, fit-for-purpose data quality, risk-adequate level of information security, and interest-oriented transparency (such as through appropriate documentation of data-related activities by data controllers) (DEK, 2019, p. 8). The DEK also point out that the extent to which individuals should be entitled to data-specific rights from participation in data generation, depends on: the nature and scope of data generation, the weight of legitimate interest in being granted data right, the weight of conflicting interests and potential compensation arrangements, the interests of the general public, and the balance of power between the involved parties (DEK, 2019, p. 9) This data perspective is complemented by the DEK's algorithms perspective, which prescribes the following principles in the use of algorithmic systems: human-centred design (i.e. prioritising human values and rights), compatibility with core societal values (e.g. democracy and fairness), sustainability, quality and performance, robustness and security, minimisation of bias and discrimination, transparent, explainable and comprehensible systems, and clear accountability structures (DEK, 2019, pp. 17-18).

[55] An effective transparency requirement would need to offer an explanation that is both feasible and useful. A feasible definition of transparency allows programmers or algorithm producers to comply with the requirement. Ideally, this means answering what were the main factors in the decision, how changing a certain factor would have changed the decision and, if applicable, what factor resulted in different decisions in two cases that look similar (Doshi-Velez & Kortz, 2017, pp. 8-9). First, a transparency requirement could focus on the input: the training, testing and operational data. Having access to this data could allow an observer to detect biases present in the dataset on which the algorithm operated or in society at large, as discussed above. Algorithms could be reviewed for their neutrality based on the accuracy and characteristics of the algorithm's inputs. A second possibility is to require transparency of the decision-making process of the algorithm. Following this approach, an explanation should permit an observer to determine how input features relate to predictions, and how influential a particular input is on the output. This presumes that the model of the algorithm can be articulated and understood by a human (Goodman & Flaxman, 2017, p. 6). An explanation of the decision model of the algorithm may become increasingly difficult as the algorithm becomes more complex. Even if an algorithm's source code were made transparent, it would only give a snapshot of its functionality. This is particularly true for adaptive, self-learning systems (Ananny & Crawford, 2016, p. 982). Finally, transparency could be required in terms of the outcomes or decisions of the algorithm. In some cases, it may be obvious that the outcome or decision reached by an algorithm is harmful. In others, analysing the outcomes may show harm, for instance in the form of (statistical) discrimination. The first and last approach to transparency may be more feasible for programmers than the approach focusing on the decision-model of the algorithm. As technology advances, more instruments may become available to quantify the degree of influence of input variables on algorithm outputs (Datta et al., 2016). Research is also underway in pursuit of rendering algorithms more amenable to ex post and ex ante inspection. (Jia & Liang, 2016) Nonetheless, generating explanations of an algorithm is a non-trivial engineering task that takes time and effort that could also be spent on other goals (Doshi-Velez & Kortz, 2017, pp. 3). The usefulness of transparency may depend on the risk associated with the decision. Regulatory transparency requirements should be context-dependent and based on risks to safety, fairness, and privacy(Wachter et al., 2017).

[56] An explanation could be provided by probing the AI system with variations of the original inputs changing only the relevant variable, to see if the outcomes are different (Doshi-Velez & Kortz, 2017, pp. 7). A simple example would be: 'You were denied a loan because your annual income was £30,000. If your income had been £45,000, you would have been offered a loan'(Wachter et al., 2018, p. 844).

[57] A special issue of Journal of Cyber Policy (volume 3, issue 3, 2018), edited by Aggarwal & Reddie (2018a), provides an overview of cybersecurity state of play several countries (US, China, Japan, Taiwan, France, UK, and Finland) and on the EU (2018), which is used below to highlight selectively some relevant aspects.

[58] Cyber risk in an Internet of Things world, Flashpoint Report, Deloitte, 2015: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/technology-media-telecommunications/us-tmt-flashpoints-cyber-risk-in-internet-of-things-world.pdf .

[59] Notable incidents related to physical and cybersecurity of energy: Reports of hackers penetrating Russian and US power networks, 2019 In March 2019, the US grid regulator NERC reportedly warned that a hacking group with suspected Russian ties was conducting reconnaissance into the networks of American electrical utilities. In June 2019, the New York Times reported that American 'code' had been deployed inside many elements of Russia's power network by US military hackers that were targeting Russian power plants. The claims were denied by President Trump and regarded with scepticism by cybersecurity experts. Cyber-attack on petrochemical plant, Saudi Arabia, August 2017 In August 2017, a sophisticated cyber-attack on a Saudi petrochemical plant was the first known attempt to manipulate an emergency shutdown system. The attack resulted in the plant shutting down, but experts warned that it had the potential to cause a serious industrial accident. Cybersecurity experts attributed the incident to a Russian government-owned laboratory. Cyber-attacks on Ukrainian power grid, 2015 and 2016 The Ukrainian grid suffered two blackouts as a result of cyber-attacks. In December 2015, hackers penetrated the computer system of a western Ukrainian power utility, and cut off the electricity to some 225 000 people. A year later, in December 2016, a cyber-attack disabled an electricity substation and left customers in parts of Kiev without power for about an hour. Both attacks were attributed to Russian hacker groups. Some security researchers suspect that the second attack was intended to cause physical damage to the components of the Ukrainian electricity grid. Metcalf sniper attack, California, 2013 In April 2013, attackers physically damaged and disabled the Metcalf substation that supplies electricity to Silicon Valley. In a well-planned night-time operation, they cut communication cables and used rifles to severely damage 17 electricity transformers, resulting in damage worth US$15 million. The attackers were not identified and their motivation is not known. Baku-Tbilisi-Ceyhan oil pipeline explosion, Turkey, 2008 The Baku-Tbilisi-Ceyhan (BTC) oil pipeline in Turkey experienced a rupture and fire in 2008. The Kurdish Workers Party claimed responsibility for the incident, but later investigations point to a cyber-attack in which the attackers accessed the control system of the pipeline via internetconnected security cameras and gained access to the industrial control systems to raise the pressure in the pipeline, causing it to rupture. North-eastern blackout, USA and Canada, 2003 Malware may have inadvertently contributed

to the 2003 blackout, which left 50 million North Americans without electricity. The blackout happened at a time when the computer worm Blaster affected a large number of computer systems, possibly impeding the timely detection of, and communication about, the initial small power outage, which cascaded to interconnected grids.

[60] Directive (EU) 2016/1148 concerning measures for a high common level of security of network and information systems across the Union.

[61] The author collected data on the decisions of 213,810 public and private companies regarding when to update their web server software and apply security patches during the years before and after the California legislation was signed. Comparisons are made between groups of companies within and outside of the jurisdiction of the California law using a difference-in-differences framework. The study shows that firms that use older server software are also more likely to suffer a successful hacking event and data breach. In addition, it found that the data breach notification law in California caused firms headquartered in that state to use web server software that was 1.8-2.8% newer.

[62] The model assumes that firms (agents) have few incentives to unilaterally report breaches. To enforce the law, regulators (principals) can introduce security audits and sanction noncompliance. The model predicts that it may be difficult to adjust the sanction level such that breach notification laws generate social benefits If disclosure costs are not negligible, a security breach notification law without security audits, regardless of the sanction level, cannot incentivize firms to report security breaches to authorities (investments made by firms are below social optimum level) a breach notification law with security audits and sanctions can incentivize firms to report breaches to authorities, regardless of accompanied disclosure costs. With such a law in place, firms face sanctions for noncompliance with reporting obligations, and indirect cost associated with information sharing. Therefore, firms conduct additional security investments to reduce their breach probabilities, and thus the number of reporting obligations.

[63] See: https://unctad.org/en/Pages/DTL/STI_and_ICTs/ICT4D-Legislation/eCom-Data-Protection-Laws.aspx

[64] See: https://www.oed.com/view/Entry/80307

[65] See: https://fxssi.com/top-10-most-valuable-companies-in-the-world

[66] See: https://solid.inrupt.com

[67] See: Deloitte blog: Block Chain. Enigma. Paradox. Opportunity (retrieved from: https://www2.deloitte.com/uk/en/pages/innovation/articles/blockchain.html , 16/10/2019).

[68] See: W3C Decentralized Identifiers (DIDs) v0.13(retrieved from https://w3c-ccg.github.io/did-spec/ , 16/10/2019).

[69] See: https://aws.amazon.com/iam/ (accessed 16/10/2019).

[70] See: https://azure.microsoft.com/en-us/services/active-directory/ (accessed 16/10/2019).

[71] See: https://hatdex.dataswift.io.

[72] Caroline Wren reported at the EDAA Summit 2019: "Since GDPR, many consumers feel more knowledgeable about online data". (See" https://digitalenlightenment.org/system/files/oliver_gray_edaa_2019_caroline_wren.pdf)

[73] See: https://www.wired.com/story/larry-sanger-declaration-of-digital-independence/

[74] A full discussion of this new productivity paradox is beyond the scope of this report, and we simply report here the summary presented in the earlier cited Briefing released by the European Parliament (2019, p. 5). In advanced economies productivity is sluggish in an age of accelerating technological progress. One traditional explanation is that of lag time: productivity gains will materialise when diffusion of AI capabilities increases, and complementary innovations are adopted. Other economists (such as Gordon cited above) consider that ICT revolution has reached maturity and show that research productivity is declining sharply with diminishing impacts on the economy. According to opposing views, AI will significantly improve human capital by offering novel ways of teaching and training the workforce. One possible explanation, then, is that productivity gain do not show due to mis-measurement.

[75] See: UK Digital Competition Expert Panel – Unlocking Digital Competition. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf

'At present, merger assessment only considers how likely a merger is to reduce competition. If a substantial lessening of competition is more likely than not to result, a merger may be blocked. Although in many situations this is a reasonable approach, it does not adequately allow the scale of any harm (or benefits) to be accounted for alongside their likelihood as they would be in economically sound cost-benefit analysis. For digital mergers, this can be a crucial gap. For example, take a large platform seeking to acquire a smaller tech company based on an attractive innovation that gives it a real chance of competing for consumers. For the sake of the example, assume that if the companies merge, there would only be a modest efficiency benefit. But if the smaller company would otherwise have become a serious and innovative competitor, the resulting competition would have generated far greater

consumer benefits. The Panel is concerned that, under the system as it stands, the CMA could only block the merger if it considered the smaller company more likely than not to be able to succeed as a competitor. This is unduly cautious. The report recommends that assessment should be able to test whether a merger is expected to be on balance beneficial or harmful, taking into account the scale of impacts as well as their likelihood. This change would move these merger decisions to a more economically rational basis, and allow big impacts with a credible and plausible prospect of occurring –critical in digital markets –to be taken properly into account. Recommended action 10: A change should be made to legislation to allow the CMA to use a 'balance of harms' approach which takes into account the scale as well as the likelihood of harm in merger cases involving potential competition and harm to innovation.'

And "The Data Ethics Commission ascribes enormous importance to a holistically conceived, sustainable and strategic economic policy that outlines effective methods of preventing not only the exodus of innovative European companies or their acquisition by third-country companies, but also an excessive dependence on third-country infrastructures (e.g. server capacities). A balance must be struck in this context between much-needed international cooperation and networking on the one hand, and on the other a resolute assumption of responsibility for sustainable security and prosperity in Europe against the backdrop of an ever-evolving global power dynamic." https://www.bmjv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN.pdf?__blob=publicationFile&v=1

[76] The FCC prohibited internet service providers, such as Comcast and Verizon, from discriminating against non-affiliated content-providers.

[77] See: https://en.wikipedia.org/wiki/Social_media_as_a_public_utility.

[78] See: one among many interviews to the Internet 'daddy' at: https://www.theregister.co.uk/2018/10/01/tim_berners_lee_solid_inrupt/ (accessed 16/10/2019).

[79] These authors work in the tradition of labour economics theorising about the creation of absent markets. In one case it has been argued that by creating or strengthening absent markets, it is possible to simultaneously address the inequality, stagnation and socio-political conflict afflicting developed countries (Posner & Weyl, 2018). Posner & Weyl, 2018 call such cases 'radical markets' because of their transformative emancipatory potential In another contribution it has been highlighted the social problems with the culture of "free" online, in which users are neither paid for their data contributions to digital services nor pay directly for the value they receive from these services (Lanier, 2013). While free data for free services is a barter, Lanier argues that the lack of targeting of incentives undermines market principles of evaluation, skews distribution of financial returns from the data economy and stops users from developing themselves into "first-class digital citizens. Within this tradition, Arrieta-Ibarra et al. (2018) explore whether and how treating the market for data like a labour market could serve as a radical market that is practical in the near term. According to their analysis, in the present situation user expectations of receiving free online service works hand in hand with the monopsony power of the technology giants to perpetuate the status quo. Dominant tech giants benefit from free / cheap data availability in what the authors call Data as Capital (DaC) equilibrium. But the total value created by data could be much larger in a Data as Labour (DaL) world, users would likely demand compensation, reducing the share of value that captured from Tech giants. In an extreme version of monopsony (usually depressing wages) in the current (DaC) equilibrium users are not even aware of the value their data daily create for Google of Facebook. So, the author claim, monopsony by the combined tech giants' model may be an important force blocking the potential productivity gains that would accrue in a DaL equilibrium. They further make their point, by differentiating different actors within the group of tech giants and between them and startups. Google and Facebook rely heavily on DaC, more than Amazon, Apple, and Microsoft that follow other business models. The latter lag behind Google and Facebook in the race to train machine learning systems with data. So, company like Microsoft might even benefit from users perceiving themselves more as producers online. Paying users as data labourers might help Microsoft and smaller companies in competing with Google and Facebook in accessing data to create AI systems. Smaller companies or start-ups could also make a difference, for many have been formed around DaL-related ideas.

[80] Many fear that artificial intelligence (AI) systems will replace human workers. Economists rightly respond that greater technological disruptions in the past, while causing shifts in employment, have largely left labour's share of income constant or even growing (Autor, 2015). Yet recent secular declines in labour's share (Karabarbounis & Neiman, 2014) belie its universal stability.

[81] Information about the methodology used to construct the index and various measurements produced in recent years can be found at: https://digital-agenda-data.eu/datasets/desi/visualizations .

[82] See: https://www.enisa.europa.eu/

[83] See: https://www.itgovernance.co.uk/nis-directive

[84] See: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679

[85] See: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019R0881&from=EN

[86] See: https://www.globalsign.com/en/blog/four-cybersecurity-regulations-you-should-know/

[87] See: http://www.ncsl.org/research/telecommunications-and-information-technology/cybersecurity-legislation-2018.aspx

[88] See: https://www.theguardian.com/commentisfree/2019/mar/09/eu-plan-facebook-google-online-copyright-law

[89] See: https://www.privacyshield.gov/welcome

[90] See: http://www.europarl.europa.eu/RegData/etudes/IDAN/2018/625151/EPRS_IDA(2018)625151_EN.pdf

[91] See: https://noyb.eu/cjeu-case/?lang=en

[92] See: http://www.europeanpapers.eu/en/e-journal/eu-us-data-transfer-safe-harbour-privacy-shield

[93] See: http://curia.europa.eu/juris/document/document.jsf?text=&docid=204046&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=6462885

[94] See: https://www.lexology.com/library/detail.aspx?g=00c338c4-358b-47ca-a4a2-39857dbcc514

[95] See: https://www.apt.int/APT-Introduction

[96] See: https://meity.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf

[97] See: https://ec.europa.eu/digital-single-market/en/trust-services-and-eid

[98] See: https://sovrin.org/faq/what-is-self-sovereign-identity/

[99] Sovrin, What is self-sovereign Identity? December 6, 2018 (retrieved from: https://sovrin.org/wp-content/uploads/2019/01/How-DIDs-Keys-Credentials-and-Agents-Work-Together-in-Sovrin-131118.pdf , 16/10/2019).

[100] See: https://sovrin.org.

[101] See W3C Decentralized Identifiers (DIDs) v0.13(retrieved from https://w3c-ccg.github.io/did-spec/ , 16/10/2019).

[102] See: http://identity.foundation (accessed 16/10/2019).

[103] See: http://identity.foundation/events/ (accessed 16/10/2019).

[104] See: www.enforcementtracker.com.

[105] We report here a few excerpts from the article's abstract. The information on cyber-crime losses mostly come from surveys that presents several shortcomings. Since losses are extremely concentrated, a representative sample of the population does not give a representative sample of the losses. Second, losses are self-reported and unverifiable numbers. The authors find evidence that most surveys are dominated by a minority of responses in the upper tail (i.e., a majority of the estimate is coming from as few as one or two responses). Finally, the fact that losses are confined to a small segment of the population magnifies the difficulties of refusal rate and small sample sizes. A single individual who claims $50,000 losses, in an N = 1000 persons survey, is all it takes to generate a $10 billion loss over the population. One unverified claim of $7,500 in phishing losses translates into $1.5 billion.

[106] This manifests itself in hesitation to buy from another country in the EU, or overlap in standards, duplication in certification schemes (ENISA, 2016)or uncertainty about how to deal with cross-border cyber incidents. Fragmentation leads to smaller markets, less economies of scale, and thereby smaller companies, making home-grown industry less competitive globally (ECIL, 2016). Small local companies are not able to invest as much as their large global competitors in brand and reputation, which are tremendously important to convey trustworthiness in the complex and often obscure field of cybersecurity.

[107] Internet of Things (IoT) Security and Privacy Recommendations, Broadband Internet Technical Advisory Group Report, 2016(BITAG, 2016). Risks of IoT are linked, among the others, to: lack of IoT supply chain experience with security and privacy; lack of incentives to develop and deploy updates after the initial sale; difficulty of secure over-the-network software updates; devices with constrained or limited hardware resources (precluding certain basic or "common-sense" security measures); devices with constrained or limited user-interfaces (which if present, may have only minimal functionality), and devices with malware inserted during the manufacturing process. Internet of Things (IoT) Security and Privacy Recommendations.

[108] Cyber risk in an Internet of Things world, Flashpoint Report, Deloitte, 2015: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/technology-media-telecommunications/us-tmt-flashpoints-cyber-risk-in-internet-of-things-world.pdf.

[109] Notable incidents related to physical and cybersecurity of energy: Reports of hackers penetrating Russian and US power networks, 2019 In March 2019, the US grid regulator NERC reportedly warned that a hacking group with suspected Russian ties was conducting reconnaissance into the networks of American electrical utilities. In June 2019, the New York Times reported that American 'code' had been deployed inside many elements of Russia's power network by US military hackers that were targeting Russian power plants. The claims were denied by President Trump and regarded with

scepticism by cybersecurity experts. Cyber-attack on petrochemical plant, Saudi Arabia, August 2017 In August 2017, a sophisticated cyber-attack on a Saudi petrochemical plant was the first known attempt to manipulate an emergency shutdown system. The attack resulted in the plant shutting down, but experts warned that it had the potential to cause a serious industrial accident. Cybersecurity experts attributed the incident to a Russian government-owned laboratory. Cyber-attacks on Ukrainian power grid, 2015 and 2016 The Ukrainian grid suffered two blackouts as a result of cyber-attacks. In December 2015, hackers penetrated the computer system of a western Ukrainian power utility, and cut off the electricity to some 225 000 people. A year later, in December 2016, a cyber-attack disabled an electricity substation and left customers in parts of Kiev without power for about an hour. Both attacks were attributed to Russian hacker groups. Some security researchers suspect that the second attack was intended to cause physical damage to the components of the Ukrainian electricity grid. Metcalf sniper attack, California, 2013 In April 2013, attackers physically damaged and disabled the Metcalf substation that supplies electricity to Silicon Valley. In a well-planned night-time operation, they cut communication cables and used rifles to severely damage 17 electricity transformers, resulting in damage worth US$15 million. The attackers were not identified and their motivation is not known. Baku-Tbilisi-Ceyhan oil pipeline explosion, Turkey, 2008 The Baku-Tbilisi-Ceyhan (BTC) oil pipeline in Turkey experienced a rupture and fire in 2008. The Kurdish Workers Party claimed responsibility for the incident, but later investigations point to a cyber-attack in which the attackers accessed the control system of the pipeline via internetconnected security cameras and gained access to the industrial control systems to raise the pressure in the pipeline, causing it to rupture. North-eastern blackout, USA and Canada, 2003 Malware may have inadvertently contributed to the 2003 blackout, which left 50 million North Americans without electricity. The blackout happened at a time when the computer worm Blaster affected a large number of computer systems, possibly impeding the timely detection of, and communication about, the initial small power outage, which cascaded to interconnected grids.

[110] Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019H0553&from=EN

[111] Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R1938&from=EN

[112] Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019R0941&from=EN

[113] Directive (EU) 2016/1148 concerning measures for a high common level of security of network and information systems across the Union.

[114] The author collected data on the decisions of 213,810 public and private companies regarding when to update their web server software and apply security patches during the years before and after the California legislation was signed. Comparisons are made between groups of companies within and outside of the jurisdiction of the California law using a difference-in-differences framework. The study shows that firms that use older server software are also more likely to suffer a successful hacking event and data breach. In addition, it found that the data breach notification law in California caused firms headquartered in that state to use web server software that was 1.8-2.8% newer.

[115] The model assumes that firms (agents) have few incentives to unilaterally report breaches. To enforce the law, regulators (principals) can introduce security audits and sanction noncompliance. The model predicts that it may be difficult to adjust the sanction level such that breach notification laws generate social benefits If disclosure costs are not negligible, a security breach notification law without security audits, regardless of the sanction level, cannot incentivize firms to report security breaches to authorities (investments made by firms are below social optimum level) a breach notification law with security audits and sanctions can incentivize firms to report breaches to authorities, regardless of accompanied disclosure costs. With such a law in place, firms face sanctions for noncompliance with reporting obligations, and indirect cost associated with information sharing. Therefore, firms conduct additional security investments to reduce their breach probabilities, and thus the number of reporting obligations.

References

Abdul-Rahman, A., & Hailes, S. (2000). Supporting trust in virtual communities. (pp. 132).

Abelson, H., Anderson, R., Bellovin, S., Benaloh, J., Blaze, M., Diffie, W., & Gilmore, J. (2015). Keys Under Doormats: Mandating Insecurity by Requiring Government Access to all Data and Communications. Journal of Cybersecurity, 1(1), 69-79.

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. Science, 347(6221), 509. doi:10.1126/science.aaa1465

Aggarwal, V., & Reddie, A. (2018a). Comparative industrial policy and cybersecurity: a framework for analysis. Journal of Cyber Policy, 3(3), 291-305.

Aggarwal, V., & Reddie, A. (2018b). Comparative Industrial Policy and Cybersecurity: The US Case. Journal of Cyber Policy, 3(3), 445-456.

Al-shamasneh, A., & Obaidellah, U. (2017). Artificial Intelligence Techniques for Cancer Detection and Classification: Review Study. European Scientific Journal, 13(3), 342-370.

Albrycht, I., & Swiatkowska, J. (2019). The Future of 5G or Quo Vadis, Europe? Krakow: he Kosciuszko Institute Policy Brief (retrieved from: https://ik.org.pl/wpcontent/uploads/ik_policy_brief_5g_eng.pdf, 21/11/2019).

Ananny, M., & Crawford, K. (2016). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. New Media & Society, 20(3), 973-989. doi:10.1177/1461444816676645.

Anderson, R., Barton, C., Bohme, R., Clayton, R., Ganan, C., Grasso, T., Moore, T. (2019June). Measuring the changing costs of cybercrime. Paper presented at the 2019 Workshop on the Economics of Information Security Boston. https://weis2019.econinfosec.org/wp-content/uploads/sites/6/2019/05/WEIS_2019_paper_25.pdf

Arrieta-Ibarra, I., Goff, L., Jiménez-Hernández, D., Lanier, J., & Weyl, E. (2018). hould We Treat Data as Labor? Moving beyond «Free». AEA Papers and Proceedings, American Economic Association, 108, 38-42.

Atali, A., Gnanasambandam, C., & Srivathsan, B. (2019). Transforming infrastructure operations for a hybrid-cloud world. McKinsey & Company (retrieved from: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/transforming-infrastructure-operations-for-a-hybrid-cloud-world# , 21/11/2019.

Atkinson, R., Castro, D., Ezell, S., McQuinn, A., & New, J. (2016). A Policymaker's Guide to Digital Infrastructure. Washington DC: Information Technology and Innovation Foundation (ITIF, retrieved from: http://www2.itif.org/2016-policymakers-guide-digital-infrastructure.pdf?_ga=2.189305762.867435639.1574195581-2032229974.1571398829

Publisher
EIT Digital
Rue Guimard 7
1040 Brussels
Belgium
www.eitdigital.eu

Contact
info@eitdigital.eu

EIT Digital is supported by EIT,
a body of the European Union

ISBN 978-91-87253-62-1