The Role of Data Regulation in Shaping AI: An Overview of Challenges and Recommendations for SMEs

Tjerk Timan, Charlotte van Oirsouw, and Marissa Hoekstra

Abstract In recent debates around the regulation of artificial intelligence, its foundations, being data, are often overlooked. In order for AI to have any success but also for it to become transparent, explainable and auditable where needed, we need to make sure the data regulation and data governance around it is of the highest quality standards in relation to the application domain. One of the challenges is that AI regulation might – and needs to – rely heavily on data regulation, yet data regulation is highly complex. This is both a strategic problem for Europe and a practical problematic: people, institutions, governments and companies might increasingly need and want data for AI, and both will affect each other technically, socially but also regulatory. At the moment, there is an enormous disconnect between regulating AI, because this happens mainly through ethical frameworks, and concrete data regulation. The role of data regulation seems to be largely ignored in the AI ethics debate, Article 22 GDPR being perhaps the only exception. In this chapter, we will provide an overview of current data regulations that serve as inroads to fill this gap.

Keywords Big data · Artificial intelligence · Data regulation · Data policy · GDPR

1 Introduction

It has been over 2 years since the introduction of the GDPR, the regulation aimed at harmonising how we treat personal data in Europe and sending out a message that leads the way. Indeed, many countries and states outside of Europe have since followed suit in proposing stronger protection on data trails we leave behind in digital and online environments. However, in addition to the GDPR, the European

T. Timan (✉) · M. Hoekstra
Strategy, Analysis & Policy Department, TNO, The Hague, The Netherlands
e-mail: tjerk.timan@tno.nl

C. van Oirsouw
Tilburg University, Tilburg, The Netherlands

E. Curry et al. (eds.), *The Elements of Big Data Value*,
https://doi.org/10.1007/978-3-030-68176-0_15

Commission (EC) has proposed and instated many other regulations and initiatives that concern data. The free flow of data agenda is meant to lead the way in making non-personal data usable across the member states and industries, whereas the Public Sector Information Directive aims to open up public sector data to improve digital services or develop new ones. Steps have also been made in digital security by harmonising cybersecurity through the NIS Directive, while on the other side law enforcement in both the sharing of data (through the e-Evidence Directive) and the specific ways in which it is allowed to treat personal data (Police Directive) has been developed. On top of this already complex set of data regulations, the new Commission has stated an ambitious agenda in which further digitisation of Europe is one of the key pillars, placing even more emphasis on getting data regulation right, especially in light of transitioning towards artificial intelligence.

Yet, however impactful and ahead-of-the-curve the regulatory landscape is, for day-to-day companies and organisations, often already part of a sector-specific set of regulations connected to data, it is not hard to see why for many states it has become difficult to know what law to comply with and how.[1] While there is no particular framework that specifically applies to (big) data, there are many frameworks that regulate certain aspects of it. In this chapter, we aim to give an overview of the current regulatory framework and recent actions undertaken by the legislator in that respect. We also address the current challenges the framework faces on the basis of insights gathered throughout the project[2] and using academic articles and interviews we held with both legal scholars and data practitioners, and multiple sessions and panels in both academic and professional conferences as a basis for this chapter.[3] One of the main challenges is to better understand the interaction between, and intersections of, data regulations and to look at how the different regulations around data interact and intersect. Many proposals have seen the light of day over the last couple of years, and, as stated, all these data-related regulations create a complex landscape that, especially for smaller companies and start-ups, is difficult to navigate. Complexity in itself should not be a concern; however, the world of data is complicated, as is regulating different facets of data. Uncertainty about data regulation and not knowing how to comply or what to comply with does leave its mark on the data-innovation landscape; guidance and clarification are key points of attention in bridging the gap between legal documents and data science practice. In this chapter, we also provide reflections and insight on recent policy debates, thereby contributing to a better understanding of the regulatory landscape and its several sub-domains. After discussing several current policy areas, we will end by providing

[1] See, for instance, the SMOOTH platform H2020 project, dedicated to helping SMEs in navigating the GDPR: https://smoothplatform.eu/about-smooth-project/.

[2] See for a recent view on the strategy by the novel Commission: http://www.bdva.eu/PositionDataStrategy.

[3] For an overview of activities, see https://www.big-data-value.eu/wp-content/uploads/2020/03/BDVe-D2.4-Annualpositionpaper-policyactionplan-2019-final.pdf, page 18.

concrete insights for SMEs on how data policy can help shape future digital innovations.

2 Framework Conditions for Big Data[4]

In previous work,[5] we have laid out a basis for looking at big data developments as an ecosystem. In doing so, we followed an approach presented by Lawrence Lessig in his influential and comprehensive publication *Code and Other Laws of Cyberspace* (Lessig, L., 2009). Lessig suggests online and offline enabling environment (or ecosystem) as the resultant of four interdependent, regulatory forces: law, markets, architecture and norms. He uses it to compare how regulation works in the real world versus the online world, in discussing the *regulability* of digital worlds, or cyberspace as it was called in 1999.[6]

In our work for the BDVe regarding data policy, we have worked along these axes in order to gather input and reflections on the development of the big data value ecosystem as the sum total of developments along these four dimensions. We have seen developments on all fronts, and via several activities throughout our interaction with the big data community. Some of the main challenges with respect to regulating data that we know from the academic debate also resonated in practice, such as the role and value of data markets and the sectoral challenges around data sharing. For example, ONYX,[7] a UK-based start-up operating in big data in the wind turbine industry, discussed their experience of vendor lock-in in the wind turbine industry and their involvement in a sector-led call for regulatory intervention from the EU. In another interview for the BDVe policy blog, Michal Gal provided an analysis of data markets and accessibility in relation to competitive advantages towards AI, for example.[8] On the level of architecture, some of the challenges concerning data sharing and 'building in' regulation can be found in the area of privacy-preserving technologies and their role in shaping the data landscape in Europe. In terms of norms and values, we want to reflect in this chapter on numerous talks and panels that delved into the topic of data ethics and data democracy. We will mainly focus on the regulatory landscape around data. In addition to norms (and values), markets and architecture, all remaining challenges in developing a competitive and value-driven Digital Single Market, there have been many legal developments in Europe that are

[4]Parts of this chapter appear in the public deliverable developed for the BDVe: https://www.big-data-value.eu/bdve-d2-4-annualpositionpaper-policyactionplan-2019-final/.

[5]See BDVe Deliverable D2.1, https://www.big-data-value.eu/bdve_-d2-1-report-on-high-level-consultation_final/.

[6]See BDVe Deliverable D2.1, p 18 and further: https://www.big-data-value.eu/bdve_-d2-1-report-on-high-level-consultation_final/.

[7]https://www.big-data-value.eu/the-big-data-challenge-insights-by-onyx-insights-into-the-wind-turbine-industry/

[8]https://www.big-data-value.eu/michals-view-on-big-data/

affecting and shaping the big data ecosystem. One of the main challenges we are facing right now is to see how, if at all, such a legal regime is up to the challenges of regulating AI and how this regulatory landscape can help start-ups in Europe develop novel services (Zillner et al. 2020).

3 The EU Landscape of Data Regulation

3.1 Data Governance Foundations

3.1.1 Data Governance and the Protection of Personal Data

Data is taking a central role in many day-to-day processes. In connecting data, ensuring interoperability is often the hardest part as the merging and connecting of databases takes a lot of curation time, as was stated by Mercè Crosas in an interview with the BDVe.[9] Therefore, it is important that data practices are arranged solidly by doing good data governance to avoid interoperability problems. In addition, data is an indispensable raw material for developing AI, and this requires a sound data infrastructure (High-Level Expert Group on Artificial Intelligence, 201) and better models on data governance. In a recent panel held during the BDV PPP Summit in June 2019 in Riga,[10] a researcher from the DigiTransScope project – a project in which an empirical deep-drive is made into current data governance models[11] – gave a definition of the concept of data governance, as follows: 'the kind of decisions made over data, who is able to make such decisions and therefore to influence the way data is accessed, controlled, used and benefited from'.[12] This definition covers a broad spectrum of stakeholders with varying interests in a big data landscape. More research is needed to find insights on the decision-making power of the different stakeholders involved so that a good balance is found between fostering economic growth and putting data to the service of public good. Concepts such as data commons (Sharon and Lucivero 2019) and data trusts have been emerging recently. Any kind of guidance should take all of these elements into account. It is important that all stakeholders are involved in the process of developing guidance, as otherwise the emergence and development of a true data economy are hampered.

In a data landscape, many different interests and stakeholders are involved. The challenging part about regulating data is the continuous conceptual flux, by which we mean that the changing meaning and social and cultural value of data is not easily captured in time or place. Yet, one can set conditions and boundaries that can aim to steer this conceptual flux and value of data for a longer foreseeable timeframe. One

[9] https://www.big-data-value.eu/the-big-data-challenge-recommendations-by-merce-crosas/

[10] See https://www.big-data-value.eu/ppp-summit-2019/.

[11] See https://ec.europa.eu/jrc/communities/en/community/digitranscope.

[12] https://ec.europa.eu/jrc/communities/en/community/digitranscope

of the most notable regulations passed recently is the General Data Protection Regulation (hereafter referred to as GDPR). With this regulation, and accompanying implementation acts in several member states, the protection of personal data is now firmly anchored within the EU. However, the distinction between personal and non-personal data has proven to be challenging to make in practice, even more so when dealing with combined datasets that are used in big data analytics. It has also recently been argued that the broad notion of personal data is not sustainable; with rapid technological developments (such as smart environments and datafication), almost all information is likely to relate to a person in purpose or in effect. This will render the GDPR a law that tries to cover an overly broad scope and it will therefore potentially lose power and relevance (Purtova 2018). In this vein, there is a need to continue developing notions and concepts around personal data and the types of data use.

For most big data analytics, privacy harm is not necessarily aimed at the individual but occurs as a result of the analytics itself because it happens on a large scale. EU regulation currently lacks in providing legal remedies for the unforeseen implications of big data analytics, as the current regime protects input data and leaves inferred data[13] out of its scope. This creates a loophole in the GDPR with respect to inferred data. As stated by the e-SIDES project recently,[14] a number of these loopholes can be addressed by court cases. The question remains as to whether and to what extent the GDPR is the suitable frame to curb such harms.

Despite many efforts to guide data workers through the meaning and bases of the GDPR and related data regulations such as the e-Privacy Regulation, such frameworks are often regarded by companies and governments as a hindrance to the uptake of innovation.[15] For instance, one of the projects within the BDV PPP found that privacy concerns prevent the deployment, operation and wider use of consumer data. This is because skills and knowledge on how to implement the requirements of data regulations are often still lacking within companies. The rapidly changing legal landscape and the consequences of potential non-compliance are therefore barriers to them in adopting big data processes. Companies have trouble making the distinction between personal and non-personal data and who owns which data. This was also reflected in a recent policy brief by TransformingTransport, which looked into many data-driven companies in the transport sector.[16] Additionally, these same companies experience trouble defining the purpose of processing beforehand, as within a big data context the purpose of processing reveals itself after processing. Mapping of data flows onto purposes of the data-driven service in

[13]Inferred data is data that stems from data analysis. The data on which this analysis is based was gathered and re-used for different purposes. Through re-use of data, the likelihood of identifiability increases.

[14]See e-SIDES, Deliverable D4.1 (2018).

[15]Big Data Value PPP: Policy4Data Policy Brief (2019), page 8. Available at https://www.big-data-value.eu/wp-content/uploads/2019/10/BDVE_Policy_Brief_read.pdf

[16]Transforming Transport, D3.13 – Policy Recommendations.

development presents difficulties, especially when having to understand which regulation 'fits' on different parts in the data lifecycle. On the other hand, sector-specific policies or best practices for sensitive personal data are perceived as assets by professionals because these give them more legal certainty, where they face big risks if they do not comply. In this sense, privacy and data protection can also be seen as an asset by companies. We feel that there is a need for governance models and best practices to show that the currently perceived dichotomy between privacy and utility is a false one (van Lieshout and Emmert 2018). Additionally, it is also important to raise awareness among companies in which scenarios concerning big data and AI are useful, and in which scenarios they are not.[17] One of the main challenges for law- and policymakers is to balance rights and establish boundaries while at the same time maximising utility (Timan and Mann 2019).

3.1.2 Coding Compliance: The Role of Privacy-Preserving Technologies in Large-Scale Analytics

One of the more formal/technical and currently also legally principled ways forward is to build in data protection from the start, via so-called privacy-by-design approaches (see, among many others, Cavoukian 2009 and Hoepman 2018). In addition to organisational measures, such as proper risk assessments and data access and storage policies, technical measures can make sure the 'human error' element in the risk assessment is covered.[18] Sometimes referred to as privacy-preserving technologies (PPTs), such technologies can help to bridge the gaps between the objectives of big data and privacy. Currently, many effective privacy-preserving technologies exist, although they are not being implemented and deployed to their full extent. PPTs are barely integrated into big data solutions, and the gap of deployment in practice is wide. The reasons for this are of a societal, legal, economic and technical nature. The uptake of privacy-preserving technologies is, however, necessary to ensure that valuable data is available for its intended purpose. In this way data is protected and can be exploited at the same time, dissolving the dichotomy of utility and privacy. To ensure this is achieved, PPTs need to be integrated throughout the entire data architecture and value chain, both vertically and horizontally. A cultural shift is needed to ensure the uptake of PPTs, as the current societal demand to protect privacy is relatively low. Raising awareness and education will be key in doing so. It is important that PPTs are not provided as an add-on but rather are incorporated into the product. There is wide agreement that the strongest parties have

[17] BigDataStack Project. Available at: https://bigdatastack.eu/

[18] Although obviously relying on technology only to solve data protection is not the way forward either, as in itself such technologies come with novel risks.

the biggest responsibilities concerning protecting privacy and the uptake of PPTs, as was also confirmed by the e-SIDES project (2018).[19]

Another point of discussion has been the anonymisation and pseudonymisation of personal data. It has also been argued that companies will be able to retain their competitive advantage due to the loophole of pseudonymised data, which allows for unfettered exploitation as long as the requirements of the GDPR are met.[20] Anonymised data needs to be fully non-identifiable and therefore risks becoming poor in the information they contain. Also, anonymisation and pseudonymisation techniques may serve as mechanisms to release data controllers/processors from certain data protection obligations related to breach-related obligations. Recent work done by the LeMO project found that anonymisation and pseudonymisation may be used as a means to comply with certain data protection rules, for instance with the accountability principle, measures that ensure the security of processing, purpose limitation and storage limitation. Pseudonymisation and anonymisation techniques can serve as a means to comply with the GDPR,[21] but at the same time, too far-reaching anonymisation of data can limit the predictability of big data analytics (Kerr 2012). However, as long as the individual remains identifiable, the GDPR remains applicable. It has been argued that, because of this, companies will be able to retain their competitive advantage by being able to unlimitedly exploit data as long as it is pseudonymised or anonymised.

3.1.3 Non-personal Data (FFoD)

In 2019, Regulation 2018/1807 on the free flow of non-personal data (FFoD) came into force, which applies to non-personal data and allows for its storage and processing throughout the EU territory without unjustified restrictions. Its objective is to ensure the free flow of data across borders, data availability for regulatory control and encouragement of the development of codes of conduct for cloud services. The FFoD is expected to eliminate the restrictions on cross-border data flows and their impacts on business, reduce costs for companies, increase competition (LeMO 2018),[22] increase the pace of innovation and improve scalability, thereby achieving economies of scale. This is all supposed to create more innovation, thereby benefiting the uptake of big data, in which the flow of non-personal data

[19]See the CJEU Google v. CNIL case (C-507/17). The CJEU decided that the right to be forgotten (RtBF, Article 17 GDPR) does not imply that operators of search engines (in this case Google) have an obligation to carry out global de-referencing if this RtBF is invoked because this would come into conflict with non-EU jurisdictions. It was also emphasised once more in this case that the right to data protection is not an absolute right.

[20]https://www.compliancejunction.com/pseudonymisation-gdpr/

[21]Specifically with the obligations of data protection by design and default, security of processing, purpose and storage limitation and data breach-related obligations.

[22]Especially in the cloud services market, start-ups increasingly rely on competitive cloud services for their own product or service.

will remain of continuing importance in addition to having solid data infrastructures. For instance, the GAIA-X Project addresses how open data plays a role in creating a data infrastructure for Europe.[23] Other more developed initiatives include European Industrial Data Spaces[24] or the MOBI network for opening up and sharing data around blockchains.[25]

The FFoD is the complementary piece of legislation to the GDPR as it applies to non-personal data. However, this distinction between the two regimes based on these concepts of personal and non-personal data is highly debated. The distinction is not easy to make in practice as datasets are likely to be mixed and consist of both personal and non-personal data. This is especially the case for big data datasets, as it is often not possible to determine which part of the set contains personal or non-personal data. This will result in it being impossible to apply each regulation to the relevant part of the dataset (LeMO 2018). In addition, as mentioned in the previous sections, these concepts are broad and subject to the dynamic nature of contextual adaptation. Whether data has economic value is not dependent on its legal classification. Hence, when facing opaque datasets, there is the risk of strategic firms on the basis of this legal classification, and they are likely to exploit the regulatory rivalry between the FFoD and the GDPR. The limitation of the FFoD to non-personal data is likely to be counterproductive to innovation, as personal data has high innovation potential as well (Graef et al. 2018). There is also further guidance needed where it concerns parallel/subsequent application of the GDPR and the FFoD, or where the two regimes undermine each other (Graef et al. 2018). Regardless of whether data is personal or non-personal, it is of major importance that it is secured. Hence, the following section addresses the EU regime on the security of data (Fig. 1).

3.1.4 Security of Data

The Cybersecurity Act (Regulation (EU) 2019/881) was adopted to set up a certification framework to ensure a common cybersecurity approach throughout the EU. The aim of this regulation is to improve the security standards of digital products and services throughout the European internal market. These schemes are currently voluntary and aimed at protecting data against accidental or unauthorised storage, processing, access, disclosure, destruction, loss or alteration. The EC will decide by 2034 whether the schemes will become mandatory.

The NIS Directive (Directive (EU) 2016/1148) puts forward security measures for networks and information systems to achieve a common level of cybersecurity

[23]Project GAIA-X, 29/10/2019. See https://www.bmwi.de/Redaktion/EN/Publikationen/Digitale-Welt/das-projekt-gaia-x-executive-summary.html.

[24]https://ec.europa.eu/digital-single-market/en/news/common-european-data-spaces-smart-manufacturing

[25]Mobility Open Blockchain Initiative (MOBI); see www.dlt.mobi/.

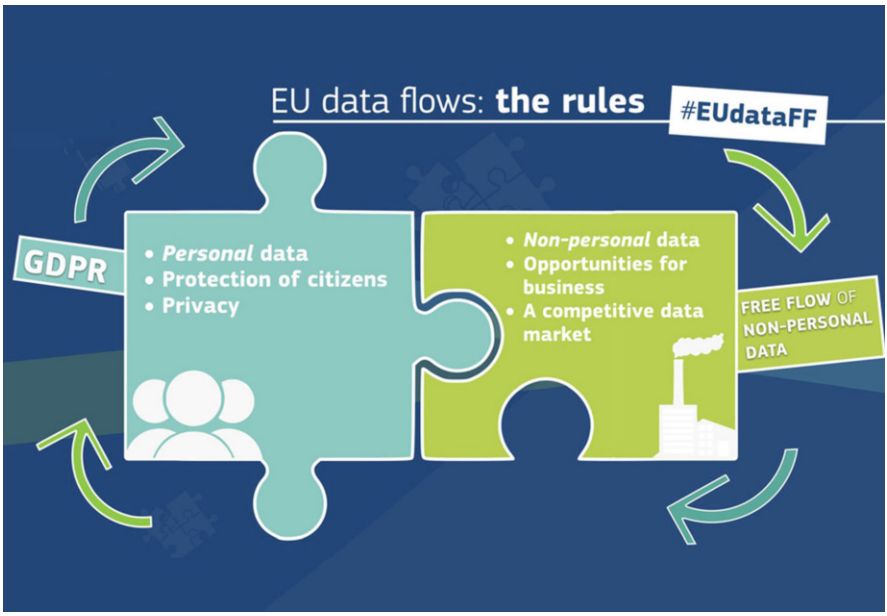

Fig. 1 The link between the GDPR and the FFoD (See https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/newsroom/eudataff_992x682px_45896.jpg) (by European Commission licensed under CC BY 4.0)

throughout the European Union to improve the functioning of the internal market. The security requirements that the NIS Directive puts forward are of both a technical and organisational nature for operators of essential services and digital service providers. If a network or information system contains personal data, then the GDPR is most likely to prevail in case of conflict between the two regimes. It has been argued that the regimes of the GDPR and the NIS Directive have to be regarded as complementary (Markopoulou et al. 2019). Cyberattacks are becoming more complex at a very high pace (Kettani and Wainwright 2019[26]). The nature of the state of play is constantly evolving, which makes it more difficult to defend against attacks. Also, it has been predicted that data analytics will be used for mitigating threats but also for developing threats (Kettani and Wainwright 2019). The companies that can offer enough cybersecurity are non-European, and the number of solutions is very limited (ECSO 2017). Due to the characteristics of the digital world, geographical boundaries are disappearing, and a report by the WRR (the Dutch Scientific Council[27]) called for attention to cybersecurity at an EU level.

Some of the characteristics of cybersecurity make tackling this challenge especially difficult; fast-paced evolvement, lack of boundaries, the fact that

[26] https://www.ecs-org.eu/documents/uploads/european-cyber-security-certification-a-meta-scheme-approach.pdf

[27] https://www.wrr.nl/

infrastructures are owned by private parties and the dependence of society on these architectures are recurring issues (ECSO 2017[28]). Currently, cyber-strategies of SMEs mainly focus on the detection of cyber risks, but these strategies should shift towards threat prevention (Bushby 2019). Just like data and robotics, AI faces all of the possible cyberthreats, and every day threats are only further evolving. Cybersecurity will also play a key role in ensuring technical robustness, resiliency and dependability. AI can be used for sophisticated automated attacks and at the same time also to provide automated protection from attacks. It is important that cybersecurity is integrated into the design of a system from the beginning so that attacks are prevented.

This section has discussed the EU regime on the security of both personal and non-personal data. Cybersecurity attacks are continually evolving and pose challenges for those involved in a data ecosystem. Keeping different types of data secure is one aspect, but successfully establishing rights upon data is another. The next section addresses the interaction between data and intellectual property rights and data ownership.

3.1.5 Intellectual Property

Due to the fact that many different players are involved in the big data lifecycle, many will try to claim rights in (part of) the datasets to protect their investment. This can be done by means of intellectual property rights. If the exercise of such a right is not done for the right reasons, this can stifle the uptake of big data and innovation. This also holds true for the cases in which an intellectual property right does not exist yet is enforced by an actor that is economically strong.

3.1.6 Public Sector Information and the Database Directive

In January 2019, an agreement was reached on the revised Public Sector Information Directive (PSI Directive). Once implemented, it will be called the Open Data and Public Sector Information Directive. The revised rules still need to be formally adopted at the time of publication of this deliverable. Public bodies hold huge amounts of data that are currently unexploited. The access and re-use of raw data that public bodies collect are valuable for the uptake of digital innovation services and better policymaking. The aim of the PSI Directive is to get rid of the barriers that currently prevent this by reducing the market entry barriers, increasing the availability of data, minimising the risk of excessive first-mover advantages and increasing the opportunities for businesses.[29] This will contribute to the growth of the EU

[28] https://www.ecs-org.eu/documents/uploads/european-cyber-security-certification-a-meta-scheme-approach.pdf

[29] EC Communication 'Towards a common European data space', SWD (2018) 125 final.

economy and the uptake of AI. The PSI Directive imposes a right to re-use data, obliges public bodies to charge the marginal cost for the data (with a limited number of exceptions), stimulates the uptake of APIs, extends the scope to data held by public undertakings, poses rules on exclusive agreements and refers to a machine-readable format when making the data available. Although open data licences are stimulated by the PSI, they can still vary widely between member states. Another challenging aspect is the commercial interests of public bodies in order to prevent distortions of competition in the relevant market. Some of the challenges that the use of public sector information faces are related to standardisation and interoperability, ensuring sufficient data quality and timely data publication, and a need for more real-time access to dynamic data. In addition, the licences to use the data can still vary, as member states are not obliged to use the standard formats. Another challenge that the PSI Directive faces is its interaction with the GDPR, either because it prevents disclosure of large parts of PSI datasets or because it creates compliance issues. The GDPR is not applicable to anonymous data. In practice, however, it is very hard for data to be truly rendered anonymous, and it cannot be excluded that data from a public dataset, combined with data from third-party sources, (indirectly) allows for identification of individuals. The interaction between the GDPR and the PSI Directive is also difficult with respect to public datasets that hold personal data, especially because of the principle of purpose limitation and the principles of data minimisation (LeMO 2018). Another challenge is the relationship of the PSI Directive with the Database Directive (DbD), as public sector bodies can prevent or restrict the re-use of the content of a database by invoking its sui generis database right. How the terms 'prevent' and 'restrict' are to be interpreted is not clear yet. Exercise of these rights bears the risk of hindering innovation. Where it concerns data portability requirements, the interaction between the DbD, PSI Directive and the GDPR is not clear either (Graef et al. 2018).

In 2018, the Database Directive (hereafter: DbD) was evaluated for the second time. The DbD protects databases by means of copyright or by means of the substantial investment that was made to create it, the sui generis right. The outcome of the evaluation was that the DbD is still relevant due to its harmonising effect. The sui generis right does not apply to machine-generated data, IT devices, big data and AI. At the time of the evaluation, a reformation of the DbD to keep pace with these developments was considered too early and disproportionate. Throughout its evaluation, one of the challenges was measuring its actual regulatory effects.

3.1.7 Copyright Reform

As part of the Digital Single Market Strategy, the EU is revising the rules on copyright to make sure that they are fit for the digital age. In 2019, the Council of Europe gave its green light to the new Copyright Directive (European Parliament, 2019). The aim is to ensure a good balance between copyright and the relevant public body objectives, such as education, research innovation and the needs of persons with disabilities. It also includes two new exceptions for Text and Data

Mining (TDM), which allows for TDM for the purpose of scientific research[30] and the opt-out clause of Article 4 New Copyright Directive. This exception will be of special importance to the uptake of AI. In a big data context, it is difficult to obtain authorisation from the copyright holder of individual data. When a work is protected by copyright, the authorisation of the rights holder is necessary in order to use the work. In a big data context, this would mean that for every individual piece of data, the authorisation needs to be obtained from the rights holder. Also, not all data in a big data context is likely to meet the originality threshold for copyright protection, though this does not exclude the data from enjoying protection under copyright. This creates uncertainties on which data is protected and which data is not, and whether a work enjoys copyright protection can only be confirmed afterwards by a court as copyright does not provide a registration system. The copyright regime is not fully harmonised throughout the EU, and a separate assessment is required on whether copyright protection is provided. This bears the potential of having a chilling effect on the uptake of EU-wide big data protection. Regarding AI-generated works of patents, it is still unclear whether, and if so to whom, the rights will be allocated. The multi-stakeholder aspect plays a role here as well, and the allocation of rights is difficult.

The manner in which intellectual property rights on data will be exercised will have a significant impact on the uptake of big data and innovation in general. This will all be shaped by the interaction between the PSI Directive, the GDPR and the new Copyright Directive. These are all instruments to establish security on data in the form of a right, as this is currently lacking.

3.1.8 Data Ownership

There is no particular framework to regulate the ownership of data. Currently, the only means to establish ownership in data or protection of data is through the provisions of the GDPR, the DbD and the Trade Secrets Protection Directive, or by contracts through contract law. Whether there should be an ownership right in data has been widely debated in recent years, as this current framework does not sufficiently or adequately respond to the needs of all the actors involved in the data value cycle. At the same time, there is consensus that a data ownership right is not desirable, as granting data ownership rights is considered to create an over-protective regime with increased data fragmentation and high transaction costs[31]. The difficulty of assigning ownership to data lies in the nature of data, because it is neither tangible nor intangible, it is limitless and non-rivalrous, and its meaning and value are not static. Data has a lifecycle of its own with many stakeholders involved. This also implies that no stakeholder will hold exclusive ownership rights over the data. The lack of a clear regulatory regime creates high levels of legal uncertainty. Ownership

[30] Article 3 Directive (EU) 2019/790 e.

[31] https://ec.europa.eu/jrc/sites/jrcsh/files/jrc104756.pdf

is currently mainly captured by contractual arrangements. This situation is far from ideal, as it creates lock-in effects and power asymmetries between parties, and is non-enforceable against third parties. However, the fact that there is no legal form of ownership does not prevent a de facto form of ownership from arising either. The rise of data bargaining markets illustrates this. The de facto ownership of data does not produce an allocation that maximises social welfare. This results in market failures, strategic behaviour by firms and high transaction costs. There is a need for policies and regulations that treat 'data as a commodity'. This requires new architectures, technologies and concepts that allow sellers and buyers of data to link and give appropriate value, context, quality and usage to data in a sense that ensures ownership and privacy where necessary.[32] In the next section, we will elaborate how this plays out in the data economy.

3.1.9 Data Economy

The digital economy is characterised by extreme returns based on scale and network effects, network externalities and the role of data in developing new and innovative services. As a result, the digital economy has strong economies of scope with large incumbent players who are difficult to dislodge. In order to realise the European Digital Single Market, we need the conditions that allow for the realisation thereof. Moreover, AI and the IoT are dependent on data; the uptake of both will be dependent on the data framework.[33]

3.1.10 Competition

There have been many developments in the field of competition law that are of importance for the regulation of big data. The legal principles of competition law stem from a time when the digital economy did not even exist yet. It has been widely debated whether the current concepts of competition law policy are sufficient tools to regulate emerging technologies or whether new tools are needed. Currently, there is still a lot of legal uncertainty concerning the practical implementation of competition law related to the data economy due to its lack of precedents. The concepts of, among others, the consumer welfare standard, the market definition and the manner in which market power is measured need to be adapted or refined in order to keep up with the digital economy (European Commission, Report - Competition policy for the Digital Era,[34]). The question of whether big tech must be broken up was often

[32]BVD PPP Summit Riga 2019, Antonis Litke, Policy4Data and DataMarketplaces ICCS/NTUA.

[33]See also the recent DataBench recommendations: https://www.databench.eu/the-project/.

[34]Available at https://ec.europa.eu/competition/publications/reports/kd0419345enn.pdf

asked in competition policy debates. Facebook is currently under investigation by the US Federal Trade Commission for potentially harming competition, and Federal Trade Commission Chairman Joe Simons has stated in an interview with Bloomberg that he is prepared to undo past mergers if this is deemed necessary to restore competition. However, there are no precedents on breaking up big tech firms, and knowledge on how to do this if considered desirable is currently lacking.[35] The aim of some of the projects that are a part of the BDVA (Zillner et al. 2017) is to make sure that we as an EU landscape become stronger through data sharing, not by aiming to create another company that becomes too powerful to fail (e.g. GAFAM). The overall aim of DataBench[36] is to investigate the current big data benchmarking tools and projects currently in operation and to identify the main gaps and provide metrics to compare the outcomes that result from those tools. The most relevant objective mentioned by many of the BDVA-related projects is to build a consensus and reach out to key industrial communities. In doing so, the project can ensure that the activity of benchmarking of big data activities is related to the actual needs and problems within different industries. Due to rules imposed by the GDPR, the new copyright rules on content monitoring and potential rules on terrorist content monitoring,[37] and realising the complexity of tasks and costs that all such regulations introduce, for the moment only large international technology companies are equipped to take up these tasks efficiently. As of this moment, there is no established consensus on how to make regulation balanced, meaning accessible and enforceable.

Over the last couple of years, several competition authorities have been active with competition law in enforcement regarding big tech. For instance, the EC has started a formal investigation into Amazon as to whether they are using sales data (which becomes available as a result of using the platform) to compete unfairly.[38] In addition, several national competition authorities have taken action to tackle market failures causing privacy issues by using instruments of competition law.[39] For example, on 7 February 2019, the German Bundeskartellamt accused Facebook of abusing its dominant position (Art. 102 TFEU) by using exploitative terms and conditions for their services. The exploitative abuse consisted of using personal data which was obtained in breach of the principles of EU data protection law. The Bundeskartellamt used the standards of EU data protection law as a qualitative parameter to examine whether Facebook had abused its dominant position. The European Data Protection Board (EDPB) also stated that where a significant merger

[35] https://www.economist.com/open-future/2019/06/06/regulating-big-tech-makes-them-stronger-so-they-need-competition-instead

[36] https://www.databench.eu/

[37] The European Parliament voted in favour of a proposal to tackle misuse of Internet hosting services for terrorist purposes in April 2019: https://www.europarl.europa.eu/news/en/press-room/20190410IPR37571/terrorist-content-online-should-be-removed-within-one-hour-says-ep.

[38] https://ec.europa.eu/commission/presscorner/detail/en/IP_19_4291

[39] For instance, the Bundeskartellamt used European data protection provisions as a standard for examining exploitative abuse: (https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2019/07_02_2019_Facebook.html).

is assessed in the technology sector, longer-term implications of the protection of economic interests, data protection and consumer rights have to be taken into account. The interaction between competition law and the GDPR is unclear, and it seems like we are experiencing a merger of the regimes, to a certain extent.

It has been considered that if substantive principles of data protection and consumer law are integrated into competition law analysis, the ability of competition authorities to tackle new forms of commercial conduct will be strengthened. If a more consistent approach in the application and enforcement of the regimes is pursued, novel rules will only be necessary where actual legal gaps occur (Graef et al. 2018). It is also been argued that, even though there are shared similarities between the regimes of competition law, consumer protection law and data protection law because they all aim to protect the welfare of individuals, competition law is not the most suitable instrument to tackle these market failures (Ohlhausen and Okuliar 2015; Manne and Sperry 2015) because each regime pursues different objectives (Wiedemann and Botta 2019). Currently, the struggle of National Competition Authorities in tackling the market failures in the digital economy creates uncertainties about how the different regimes (of competition and data protection) interact, and this creates legal uncertainty for firms.

Even though competition authorities have been prominent players in the regulation of data, the lack of precedent creates much uncertainty for companies. The next section will discuss how data sharing and access, interoperability and standards play a role in this.

3.1.11 Data Sharing and Accessibility

Data is a key resource for economic growth and societal progress, but its full potential cannot be reaped when it remains analysed in silos (EC COM/2017/09). More industries are becoming digitised and will be more reliant on data as an input factor. There is a need for a structure within the data market that allows for more collaboration between parties with respect to data. Data access, interoperability and portability are of major importance to foster this desired collaboration. In this respect, data integrity and standardisation are reoccurring issues. Accessibility and re-use of data are becoming more common in several industries, and sector-specific interpretations of the concept could have spill-over effects across the data economy. There is a need for governance and regulation to support collaborative practices. Currently, data flows are captured by data-sharing agreements.

The complexity of data flows, due to the number of involved actors and the different sources and algorithms used, makes these issues complicated for the parties involved. The terms in data-sharing agreements are often rather restrictive in the sense that only limited access is provided. This is not ideal, as restriction in one part of the value chain can have an effect on other parts of the data cycle. Access to data is mainly restricted because of commercial considerations. An interviewee suggested that the main reason that full data access is restricted is that it allows the holder of the entire dataset to control its position on the relevant market, not because of the

potential value that lies in the dataset.[40] Parties are often not aware of the importance of having full access to the data that their assets produce, resulting in the acceptance of unfavourable contractual clauses. The interviewee also suggested, however, that the real value creation does not lie in the data itself, but in the manner in which it is processed, for instance by combining and matchmaking datasets. In addition, there is a lack of certainty regarding liability issues in data-sharing agreements. Data-sharing obligations are currently being adopted in certain sectors and industries, for instance in the transport sector (LeMO 2018[41]), though due to the absence of a comprehensive legal framework, these still face numerous limitations. In some cases, the imposition of a data-sharing obligation might not be necessary as data plays a different role in different market sectors. It is worthwhile to monitor how the conditions imposed by the PSI Directive on re-use and access for public sector bodies play out in practice to see whether this could also provide a solution in the private sector (LeMO 2018).

The right to data portability of Article 20 GDPR (RtDP) is a mechanism that can facilitate the sharing and re-use of data, but regarding its scope and meaning, many areas are still unresolved. For instance, a data transfer may be required by the data subject where this is considered 'technically feasible', though what circumstances are considered to be 'technically feasible' by the legislator are not clear. In addition, there is no clarity on whether the RtDP also applies to real-time streams, as it was mainly envisaged in a static setting. There is also a strong need to consider the relationship between the right to data portability and IP rights, as it is not clear to what extent companies are able to invoke their IP rights on datasets that hold data about data subjects.[42] The interpretation of these concepts will make a big difference with respect to competition, as the right to data portability is the main means for data subjects to assay the counter-offers of the competitors for the services they use without the risk of losing their data. However, if competition law has to enforce the implementation and enforcement of interoperability standards that ensure portability, it will be overburdened in the long run.

The sharing and re-use of data require that effective standards are set across the relevant industry. Currently, the standardisation process is left to the market, but the efficient standards are still lacking, and this slows down data flows. Setting efficient standards will smoothen the process of data sharing and therefore also encourage it. Each market has its own dynamics, so the significance of data and data access will also be market dependent. In the standardisation process, it needs to be taken into account that a standard in one market might not work in another. Guidance on the creation of standards is needed to provide more legal certainty, because if this process is left to the market alone, this can result in market failures or standards that raise rivals' costs. The role of experts in the standardisation process is crucial, as

[40]This point has been made in an interview with ONYX InSight. See https://www.big-data-value.eu/the-big-data-challenge-insights-by-onyx-insights-into-the-wind-turbine-industry/.

[41]https://lemo-h2020.eu/

[42]See https://www.big-data-value.eu/spill-overs-in-data-governance/.

a deep understanding of the technology will lead to better standards. In addition, due to the multidisciplinary nature of many emerging technologies, the regulator should not address the issue through silos of law but have a holistic approach and work in regulatory teams consisting of regulatory experts that have knowledge of the fields relevant in setting the standard.[43]

Data access, interoperability, sharing and standards are important enabling factors for the data economy. The manner in which the data economy will be shaped will have an impact on commerce, consumers and their online privacy. The next section discusses these three points.

3.1.12 Consumers, e-Commerce and e-Privacy

In January 2018, the Payment Services Directive (PSD2) became applicable. This Directive was expected to make electronic payments cheaper, easier and safer. On 11 April 2018, the EC adopted the 'New Deal for Consumers' package. This proposal provides for more transparency in online marketplaces and extends the protection of consumers in respect of digital services, as they do not pay with money but with their personal data. The new geo-blocking regulation that entered into force will prohibit the automatic redirecting and blocking of access, the imposition of different general conditions to goods and services, and payment transactions based on consumer nationality. Furthermore, the EU has been working on the revision of the Civil Procedure Code regulation on consumer protection (Regulation (EC) 2017/2394), which entered into force on 17 January 2020. The new rules for VAT for the online sale of goods and services will enter into force in 2021. The Digital Services Act is a piece of legislation which is planned to tear up the 20-year-old e-Commerce Directive; it also targets Internet Service Providers (ISPs) and cloud services. It is likely to contain rules on transparency for political advertising and force big tech platforms to subject their algorithms to regulatory scrutiny (Khan and Murgia 2019). In the Communication on online platforms (Communication 2016 288[44]), the EC formulated principles for online platforms. These are about creating a level playing field, responsible behaviour that protects core values, transparency and fairness for maintaining user trust, and safeguarding innovation and open and non-discriminatory markets within a data-driven economy. Following this Communication, on 12 June 2020, the Regulation on platform-to-business relations (Regulation (EU) 2019/1150) was adjusted and is now applicable. The objective is to ensure a fair, predictable, sustainable and trusted online business environment within the internal market. Due to the scale and effects of platforms, this measure is taken at EU level instead of member state level. It applies to online intermediation services, business users and corporate website users, and it applies as soon as the business user or the corporate website user has an establishment within the EU. It sets

[43] https://www.big-data-value.eu/michals-view-on-big-data/

[44] https://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/1-2016-288-EN-F1-1.PDF

requirements for the terms and conditions, imposes transparency requirements and offers redress opportunities.

The European Data Protection Supervisor has stressed the urgency for new e-privacy laws (Zanfir-Fortuna 2018), and since the publication of the previous deliverable in 2017, the e-Privacy Directive has been under review. Several governments and institutions have expressed their opinion on its current new draft. For example, the German government has stated that they do not support the current draft version as it does not achieve the objective of guaranteeing a higher level of protection than the GDPR,[45] and the Dutch Data Protection Authority has stated that cookie walls do not comply with EU data protection laws.[46] Furthermore, in October 2019, the Court of Justice of the European Union (CJEU) gave its decision in the Planet49 case (C-673/17, ECLI:EU:C:2019:801) and stated that the consent which a website user must give for the storage of and access to cookies is not valid when this consent is given by means of a pre-ticked checkbox. In addition, information that the service provider gives to the user must include the duration of the operation of cookies and whether or not third parties may have access to these cookies. This judgement will have a significant impact on the field of e-privacy and on big data in general as well, as a lot of the data that 'forms part of big data' was gathered and processed on the basis of pre-clicked consent-box cookies. Thus, this judgement will change how data should be processed from now on.[47] In extension thereof, the case Orange Romania (C-61/19) is currently pending at the CJEU for a preliminary ruling on what conditions must be fulfilled in order for consent to be freely given.

4 Conclusions

In this chapter, some of the main challenges and developments were addressed concerning the regulatory developments in (big) data. Where across the board the main development in Europe would be the GDPR, we have tried to show that many other regulatory reforms have taken place over the last years – regulations that, similar to the GDPR, affect the data ecosystem. In areas such as competition, IP, data retention, geographical data 'sovereignty' and accessibility, the shaping of data markets, cybersecurity and tensions between public and private data, among others, we have aimed to summarise the plurality of regulatory reform and, where possible,

[45] https://www.technologylawdispatch.com/2019/08/privacy-data-protection/update-on-eprivacy-regulation-current-draft-does-not-guarantee-high-level-of-protection-and-cannot-be-supported-german-government-states/

[46] https://autoriteitpersoonsgegevens.nl/nl/nieuws/ap-veel-websites-vragen-op-onjuiste-wijze-toestemming-voor-plaatsen-tracking-cookies (in Dutch).

[47] See https://pdpecho.com/2019/10/03/planet49-cjeu-judgment-brings-some-cookie-consent-certainty-to-planet-online-tracking/.

how they intersect or interplay. Moreover, aside from the novel proposals and developments from the regulator, we have also seen the first effects of the GDPR coming into force in the form of first fines handed out to companies and local governments.[48] and we have seen other major court decisions that will have a profound effect on the data landscape (e.g. the Planet49[49] decision on cookie regulation).

To summarise our findings, the challenging aspect of regulating data is its changing nature, meaning and value. There is a need for more research on how to shape data governance models and how to implement them. The GDPR is often regarded by companies as a hindrance to innovation, but privacy and data protection can also be regarded as an asset. The implementation of privacy-preserving technologies (PPTs) can help to bridge this gap, but a gap exists in terms of their implementation in practice. Anonymisation and pseudonymisation are often used as a means to comply with the GDPR. In practice, datasets are likely to consist of both personal and non-personal data. This creates difficulties in the application of both the GDPR and the FFoD to big data. The regulatory rivalry of the GDPR and FFoD is likely to be exploited. Clarity on parallel or subsequent application of the GDPR and the FFoD is needed. Regarding the security of data, several strategies have been implemented at EU level to tackle cybersecurity issues. The nature of cybersecurity challenges makes it difficult to tackle them. Looking ahead, cybersecurity will play a key role in the development of AI and as such is a key condition for AI to shape. Another key condition for big data and AI is the use of public sector data. Use of public sector information will be challenging due to the obstacles related to data governance, for instance ensuring interoperability. Where public sector information holds personal data, the PSI will face difficulties in the interaction with the GDPR. Public sector bodies can prevent the re-use of the content of a database by invoking the sui generis database right of the Database Directive. The interaction between the PSI Directive, the GDPR and the Database Directive is not clear yet where it regards data portability requirements. In a big data context, it remains uncertain which pieces of data enjoy copyright protection under the current regime, and, connected to this, the allocation of rights for AI-generated works remains unclear.

4.1 Recommendations for SMEs and Start-Ups

The previous section gave an overview of the current regulatory landscape. It addressed the foundations of data governance, intellectual property and the data economy, thereby also revealing the uncertainties and unclarities that these frameworks face in the light of big data. In this section, we will present some concrete

[48] See, for instance, enforcementtracker.com where all fines under the GDRP are being tracked.

[49] See C-673/17, ECLI:EU:C:2019:801.

insights and recommendations for SMEs and start-ups in how data policy can help shape future digital innovations.[50]

4.1.1 Potential of Privacy-Preserving Technologies

PPTs can help SMEs to bridge the gaps between the objectives of big data and privacy.[51] The GDPR is often regarded by companies as a hindrance to innovation, but privacy and data protection can also be regarded as an asset. PPTs have great potential for SMEs, because SMEs can use them to ensure that valuable data is available for its intended purpose and that their data is protected at the same time, dissolving the dichotomy of utility and privacy. However, it is important that PPTs are not provided as an add-on but are incorporated into the product.

4.1.2 Distinction Between Personal and Non-personal Data

Anonymisation and pseudonymisation of data are often used as a means to comply with the GDPR. However, SMEs should be aware that in practice, datasets are likely to consist of both personal and non-personal data. This creates difficulties in the application of both the GDPR and the FFoD to big data. As a result, the regulatory rivalry of the GDPR and FFoD is likely to be exploited.

4.1.3 Data Security

At the moment, SMEs mainly focus their cyber-strategies on the detection of cyber risks. However, it is of major importance that cyber-strategies of companies also focus on cyber defence. For example, if cybersecurity is integrated into the design of a system from the beginning, attacks can be prevented. SMEs should therefore shift their focus from the detection of cyber risks to threat prevention in order to keep their data fully secure.

4.1.4 Intellectual Property and Ownership of Data

Due to the nature of data, it is difficult to assign ownership. Data is neither tangible nor intangible, it is limitless and non-rivalrous, and its meaning and value are not static. Currently there is no particular framework to regulate the ownership of data.

[50] See also https://www.big-data-value.eu/the-big-data-challenge-3-takeaways-for-smes-and-startups-on-data-sharing-2/.

[51] See, for example, the SODA project, which enables multiparty computation (MPC) techniques for privacy-preserving data processing (https://www.soda-project.eu/).

The only means to establish ownership of data or protection of data is through the provisions of the GDPR, the DbD and the Trade Secrets Protection Directive, or through contracts by means of general contract law.

4.1.5 Use of Consumer Data: Importance of Transparency and Informed Consent

Consumer data plays an important role in the big data landscape. When companies collect consumer data, it is important that they are transparent towards consumers about what type of data they are collecting, and that consumers give informed consent. The previously mentioned Planet49[52] decision on cookie regulation is a case in point. The way forward for EU data companies aiming to use consumer data is to step from behind the curtain and be open about data practices and underlying algorithms. Taking citizens and consumers with them on a data journey, and truly developing inclusive digital services that take the necessary organisational and technical safeguards seriously from the start (and not after the fact), might seem to many business developers like the long and winding (and far more expensive) road. However, from the insights we have gathered from policymakers, data scientists and data workers, we strongly recommend looking at data policy not as a compliance-checklist exercise but as a strong attempt to create a human rights-based competitive and fair Digital Single Market.

Acknowledgements The research leading to these results received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 732630 (BDVe).

References

Botta, M., & Wiedemann, K. (2019). The interaction of EU competition, consumer, and data protection law in the digital economy: The regulatory dilemma in the Facebook odyssey. *The Antitrust Bulletin, 64*(3), 428–446.

Bushby, A. (2019). How deception can change cyber security defences. *Computer Fraud & Security, 2019*(1), 12–14.

Cavoukian, A. (2009). Privacy by design: The 7 foundational principles. *Information and privacy commissioner of Ontario, Canada*, *5*.

European Parliament. (2019). *Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC*.

Graef, I., Husovec, M., & Purtova, N. (2018). Data portability and data control: lessons for an emerging concept in EU law. German Law Journal, 19(6), 1359–1398.

Hoepman, J. H. (2018). Privacy design strategies (the little blue book).

Kerr, O. S. (2012). The mosaic theory of the fourth amendment. *Michigan Law Review, 111*(3), 45.

[52]See C-673/17, ECLI:EU:C:2019:801.

Kettani, H., & Wainwright, P. (2019, March). On the top threats to cyber systems. In 2019 IEEE 2nd International Conference on Information and Computer Technologies (ICICT) (pp. 175–179). IEEE.

Lessig, L. (2009). *Code: And other laws of cyberspace*. ReadHowYouWant.com.

Manne, G. A., & Sperry, R. B. (2015). *The problems and perils of bootstrapping privacy and data into an antitrust framework, 2 CPI ANTITRUST CHRON. 1 (2015); Giuseppe colangelo & mariateresa maggiolino, data protection in attention markets: Protecting privacy through competition?, 8 J. EUR. C.*

Markopoulou, D., Papakonstantinou, V., & de Hert, P. (2019). The new EU cybersecurity framework: The NIS Directive, ENISA's role and the General Data Protection Regula-tion. Computer Law & Security Review, 35(6), 105336.

Purtova, N. (2018). The law of everything. Broad concept of personal data and future of EU data protection law. Law, Innovation and Technology, 10(1), 40–81.

Sharon, T., & Lucivero, F. (2019). Introduction to the special theme: The expansion of the health data ecosystem – Rethinking data ethics and governance. *Big Data & Society, 6*, 205395171985296. https://doi.org/10.1177/2053951719852969

Timan, T., & Mann, Z. Á. (2019). *Data protection in the era of artificial intelligence. Trends, existing solutions and recommendations for privacy-preserving technologies*. BDVA.

van Lieshout, M., & Emmert, S. (2018). RESPECT4U -- Privacy as innovation opportunity. In M. Medina, A. Mitrakas, K. Rannenberg, E. Schweighofer, & N. Tsouroulas (Eds.), *Privacy technologies and policy* (pp. 43–60). Cham: Springer International Publishing.

Wiedemann, K., & Botta, M. (2019). The interaction of EU competition, consumer and data protection law in the digital economy: The regulatory dilemma in the facebook odyssey. *The Antitrust Bulletin, 64*(3), 428–446.

Zillner, S., Curry, E., Metzger, A., Auer, S., & Seidl, R. (Eds.). (2017). *European Big Data Value Strategic Research & Innovation Agenda*. Retrieved from Big Data Value Association website: www.bdva.eu

Zillner, S., Bisset, D.,Milano,M., Curry, E., Hahn, T., Lafrenz, R., et al. (2020). Strategic research, innovation and deployment agenda - AI, data and robotics partnership. Third Release (3rd). Brussels: BDVA, euRobotics, ELLIS, EurAI and CLAIRE.