THE POLITICAL ECONOMY OF IOT DATA – THE STATUS QUO AND THE IMPACT OF THE DATA ACT

by Gospodin Tonev

LLM Final Thesis
SUPERVISOR: Tommaso Soave
Central European University Private University
Quellenstrasse 51-55, 1100 Vienna
Austria

June, 2023

TABLE OF CONTENTS

Executive Summary (or Abstract) 3

Introduction 4

I. Data: definition, types and choracteristics 9

I.1. Defining Data 9

I.1.2. The Three Layers 11

I.2. Types of Data 12

I.2.1. Raw & Generated Data 12

I.2.2. Co-generated Data 15

I.2.3. Personal and Non-personal Data 16

I.3. Data as a Resource 19

I.3.1. (Non-)rival and (Non-)excludable resource 20

II. Distribution of Data Wealth 25

II.1. Data Access 26

III. The Impact of the Data Act on the Status Quo 39

III.1. Context 39

III.2. Main Objectives, Principles and Instruments of the Data Act 40

III.3. Scope Ratione Materiae of the Data Act 42

III.4. Status of the IoT Users, Data Holders and Third Parties 46

III.4.1 *Status of the IoT Users* 46

III.*4.2. Status of the Data Holders* 48

III.4.*3. Status of the Third Parties* 49

Conclusion 51

Bibliography 52

EXECUTIVE SUMMARY (OR ABSTRACT)

In view of the importance of data within the Data economy the key political economy question of the current digital era is how the governance of this resource should be designed i.e., who has control over data, who can use it, and who can benefit from its value?

The question is complicated and requires understanding of the unique characteristics of the object of regulation. However, the complexity of the subject matter cannot serve as an excuse for non-regulation. On the contrary, through intensify discussions involving experts with different background – economists, lawyers and data scientist, it should be found what is the most effective and efficient model for distribution of the wealth of data and how it can be implemented in practice.

The European Commission has already taken several steps towards the regulation of the data governance and this should be encouraged and supported. The outlined by the Commission objectives are sensible. However, the development path in this field is not straightforward and this is obvious from the debates which surround virtually each legislative proposal in the area.

In 2022 the Commission announced its proposal for Data Act. This instrument has the potential to reshape the political economy of IoT data and to overcome (some of) the market failures which have been identified so far.

The current work aims to identify some possible weaknesses of the proposal on conceptual level and to contribute to the debate over the effective distribution of the value of data. For this purpose, the paper presents term ‘data’ and analyses the economic characteristics of data as a resource. Further, in order to identify possible opportunities and risks the paper explores the existing political economy of data and the two main conceptual extremes - free and unrestricted access to data and full enclosure of data.

INTRODUCTION

The economic importance of data resources is now widely recognized among scholars, practitioners and politicians – the asset of data is believed to be 'at the heart of the digital economy'[1]. This is because data economy not only means an economy within which the access to data is a factor in the control of economic processes (data-driven or data-controlled economy), but also an economy with data as goods[2]. At the same time, it is noteworthy that, within this economy, emerging because of the 'forth industrial revolution' (so-called Industry 4.0, IoT applied to manufacturing), all companies become to a greater or lesser extent data producers, even if their corporate purpose is aimed at completely different business[3]. These two facts indicate two of the main characteristics of data economy with regard to data – on the one hand, the process of data generation is amplified i.e., the number of 'data producers' increases to become (virtually) equal to the number of the natural and legal persons and simultaneously 'the amount of data generated increases exponentially'[4], on the other hand, data are recognized as goods (resources), and what is more, the data has become an essential resource for economic growth, job creation and societal progress[5]. Therefore, the access to large-scale datasets is nowadays recognized as sine qua non for any economic actor to reap the

[1] Parminder Jeet Singh and Jai Vipra, 'Economic Rights Over Data: A Framework for Community Data Ownership' (2019) 62 Development 53, 54.
[2] Herbert Zech, 'Data as a Tradeable Commodity' in Alberto De Franceschi (ed), *European Contract Law and the Digital Single Market: The Implications of the Digital Revolution* (Intersentia 2016) 56 <https://www.cambridge.org/core/books/european-contract-law-and-the-digital-single-market/data-as-a-tradeable-commodity/9D9E07D5B5E3C86C1B2E04BF8C01D15B> accessed 20 May 2023.
[3] ibid 59.
[4] Thomas Tombal, 'The Rationale for Compulsory B2B Data Sharing and Its Underlying Balancing Exercises' (1 December 2021) 1 <https://papers.ssrn.com/abstract=3984873> accessed 21 May 2023.
[5] EC Communication 'Building a European Data Economy', COM/2017/09 final.

benefits of data-driven innovation[6]. It is worth mentioning also that the value of the data market is expected to reach between 432 and 827 billion euros by 2025[7].

Under the circumstances, the key political economy question of the digital era is how the governance of data should be designed i.e., who has control over data, who can use it, and who can benefit from its value[8]? This problem is crucial because the answers to several central for the economy questions - how much corporations will profiteer at consumers' expense, whether small economic actors will survive, and if so on what terms, and which countries will move up the global ranks and which face colonization-like conditions, would all be determined by how access to and control of various important data get configured[9].

Paradoxically, despite the enormous volume of data generated and data's ground-braking potential, it is 'no nearer to being a widely available resource'[10], which motivates many legal scholars and eventually the EC to recognize the importance of adopting legal rules aiming at enhance access to data[11].

At the same time, it is clear that 'data transaction business cannot be expected to sit idle and wait until the law has figured out its conundrums'[12]. Non-legal rules invoked by market forces will distribute wealth and resources across society in the absence of legal framework. However, the problem is that 'in the context of the free flow of data there is no reason to assume per se

[6] Tommaso Fia, 'An Alternative to Data Ownership: Managing Access to Non-Personal Data through the Commons' (2021) 21 Global Jurist 181, 181.
[7] The European Data Market Monitoring Tool Key Facts & Figures, First Policy Conclusions, Data Landscape and Quantified Srories D2.9 Final Study Report, available at D29_EDM_Final_study_report_16062020_IDC_dJKdnBNy1THo2zUYc9jaJAOvk_93687.pdf, 40.
[8] Wolfgang Kerber, 'Governance of IoT Data: Why the EU Data Act Will Not Fulfill Its Objectives (Second Version)' (18 July 2022) 120 <https://papers.ssrn.com/abstract=4080436> accessed 19 April 2023; Singh and Vipra (n 1) 56.
[9] Singh and Vipra (n 1) 53.
[10] Fia (n 6) 182.
[11] ibid 184.
[12] 'Caught in the Acts: Framing Mandatory Data Access Transactions Under the Data Act, Further EU Digital Regulation Acts, and Competition Law by Peter Georg Picht :: SSRN' 9 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4076842> accessed 19 April 2023.

that market forces will efficiently allocate non-personal data'[13]. On the contrary, the lion's share of wealth now flows to those who aggregate and analyze data[14] - technological, behavioral, and legal barriers, utilized by data holders, restrict the access to large-scale datasets[15] and create factual situation identified as 'de facto data ownership'. This gives some authors a reason to expressively claim that like the conquerors 'the large players of surveillance capitalism 'claim human experience as raw material free for the taking', and purport to possess a 'right to own the behavioral data derived from human experience'—declarations that effectively render the age of surveillance capitalism 'an age of conquest'[16].

The problem of non-personal data that are not reused and shared enough (especially for innovation) seems to be the starting point of the European Data Economy initiative[17].

To achieve four main objectives within this initiative, namely: '1) Empowerment of consumers and businesses to have more control over the use of their IoT data and to benefit from more, better and cheaper products and services on secondary markets (also through more competition). 2) Making more data available to businesses, especially for more innovation (unlocking the wealth of existing data). 3) Fairness in the allocation of value from data among actors in the data economy. 4) Preserving incentives to invest in ways of generating value from data'[18], in 2022 the EC announced its proposal for a new regulation, entitled 'Data Act'[19].

[13] Laura Somaini, 'Regulating the Dynamic Concept of Non-Personal Data in the EU: From Ownership to Portability' (2020) 6 European Data Protection Law Review (EDPL) 84, 86.
[14] Lanier, J. (2014). Who owns the future? New York: Simon & Schuster, in Patrik Hummel, Matthias Braun and Peter Dabrock, 'Own Data? Ethical Reflections on Data Ownership' (2021) 34 Philosophy & Technology 545, 556.
[15] Fia (n 6) 185.
[16] Zuboff, S. (2019). The age of surveillance capitalism. London: Profile., in Hummel, Braun and Dabrock (n 14) 561.
[17] Inge Graef, Raphael Gellert and Martin Husovec, 'Towards a Holistic Regulatory Approach for the European Data Economy: Why the Illusive Notion of Non-Personal Data Is Counterproductive to Data Innovation' (27 September 2018) 14 <https://papers.ssrn.com/abstract=3256189> accessed 8 June 2023.
[18] Kerber (n 8) 122.
[19] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on harmonized rules on fair access to and use of data (Data Act), COM(2022) 68 final.

Meanwhile, alternative concepts of assigning the bundle of rights on non-personal data and allocation of the wealth of data have been proposed in the doctrine.

In the light of the foregoing and in accordance with the understanding that the distribution of wealth and resources across society because of regulatory intervention has implications for law and relates to legal framework[20], the present work attempts to answer what is the 'political economy of data' underlying the EU DA? How will the Act change the status quo? And will it be able to achieve its stated objectives?

The answers of these questions are important because, on the one hand, as it has been stated above, the central political economy question within Industry 4.0 is how the economic rights to data are allocated, and on the other hand, the proposed DA, which as main element of the EU Data Strategy is expected to response this inquiry, seems to be relatively clear with regards to its aims, but quite ambivalent when it comes to its rules. Therefore, there is need, in the first time, for clarification of the concept adopted and further for an assessment of its potential practical implications.

Should there be a clarification that the current paper focuses on the political economy on non-personal data. This is not because personal data do not have economic value and importance for the economy, but because the access to them is governed by special rules which do not aim to provide fair and effective distribution of resources, but rather to protect specific legitimate interests of the individuals – 'the fundamental right to the protection of personal data safeguards the personality of data subjects, not their property'[21].

[20] Hummel, Braun and Dabrock (n 14) 551.

[21] S. Rodotà, 'Data Protection as a Fundamental Right', in Florent Thouvenin and Aurelia Tamò-Larrieux, 'Data Ownership and Data Access Rights: Meaningful Tools for Promoting the European Digital Single Market?' in Mira Burri (ed), *Big Data and Global Trade Law* (Cambridge University Press 2021) 322 <https://www.cambridge.org/core/books/big-data-and-global-trade-law/data-ownership-and-data-access-rights/BC314C63C58A09C4B9C5D55894FE68C6> accessed 11 June 2023.

Also, the accent is on private sector data[22], which are arguably 'the most numerous and possibly valuable ones'[23].

The first chapter of the current study examines the term 'data' and the concept of data. Furthermore, it presents main characteristics of data as a resource.

The second chapter presents the two main conceptual extremes for allocation of the rights over data: (i) on the one hand, free and unrestricted access to data; (ii) on the other hand, full enclosure of all data. Also, the status quo before the DA is analyzed.

The third chapter is devoted to the DA itself and the architecture of the political economy of data enshrined in the proposal.

[22] The access and reuse of public sector data are governed by Directive (EU) 2019/1024 on open data and the re-use of public sector information.

[23] Charlotte Ducuing and others, 'White Paper on the Data Act Proposal' (26 October 2022) 15 <https://papers.ssrn.com/abstract=4259428> accessed 16 June 2023.

I. DATA: DEFINITION, TYPES AND CHORACTERISTICS

To provide an in-depth analysis of the political economy of non-personal data, which is needed for the aims of the DA to be understood and its toolkit to be evaluated, the research starts with some clarifications regarding the term 'data' and explanation of the main characteristics of data as a resource. The various possible meanings of the term 'data' should be understood in order the nuances of the legal definition to be identified – the way in which the legislator define data presupposes the scope of regulation.

I.1. Defining Data

There are attempts the term 'data' to be determined from different perspectives. From technical point of view data is 'reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing', data 'can be processed by humans or by automatic means'[24]. The ALI-ELI Principles provide the following legal definition of data: 'information recorded in any machine-readable format suitable for automated processing, stored in any medium or as it is being transmitted'[25]. It is crucial that the second definition put an accent on the fact that data must be recorded in any machine-readable format and stored (or be in state of transmission), because this is an important first step towards making data fit for being a (potential) object of legal rights[26]. Similarly, to air and water, data are omnipresent - the recording of data demarcates it similarly to the process of bottling of water after which there is sufficiently defined, classified and specified object of property rights; furthermore,

[24] ISO/IEC 2382:2015(en) Information technology — Vocabulary 2121272.
[25] Seong-Yeob Lee, Shindong Jung and Won Jae Hwang, 'ALI-ELI PRINCIPLES FOR A DATA ECONOMY - DATA TRANSACTIONS AND DATA RIGHTS -' (2022) 107 Korea Law Review 493., Principle 3(1)(a)
[26] Sjef van Erp and Koen Swinnen, 'The Legal Status of Co-Generated Data: With Particular Focus on the ALI-ELI Principles for a Data Economy and the Rules on Accession, Commingling and Specification' (2022) 2022 Technology and Regulation 61, 63.

storing is a precondition for putting a good in the course of trade. It can be concluded that 'specification functions as the container that captures data and, by doing so, makes it legally manageable'[27].

In conformity with the ALI-ELI Principles the DA proposal defines data as 'any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audio-visual recording' (Art. 2). Both the Digital Markets Act (DMA)[28] and the propose Data Governance Act (DGA)[29] provide same definition and thus it represents a common feature of the emerging EU legal framework on data sharing. This definition, however, does not reveal much about data as a legal object or as an economic good which is needed when one talks about data as an object of contracts, (property) rights or protection.

Some legal scholars reasonably argue that '[f]rom a private law viewpoint any definition of data is legally almost meaningless, as are definitions of the air, rain and sunshine'[30], yet they hold that the lack of proper demarcation of data impedes the creation of a legal framework for governing data transactions[31]. Therefore, before coming to the substance of the present paper, I would like to start with some clarifications regarding the term 'data' and some elucidations of the main features of data as resource.

[27] ibid.
[28] Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act) (Text with EEA relevance) PE/17/2022/REV/1, OJ L 265, 12.10.2022, Art. 2(24)
[29] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European data governance (Data Governance Act), COM/2020/767 final, Art. 2(1)
[30] Erp and Swinnen (n 26) 61.
[31] ibid.

I.1.2. The Three Layers

It must be clear that the term 'data' is multifaceted. Drawing from semiotics, some authors[32] distinguish between the semantic level of information (meaning), the syntactic level of information (signs and their relation with each other; data as code), and the structural level of information (physical embodiment of information; communication channel) which leads to the distinction between the content layer, code layer and physical layer.

Despite the fact that the three levels are connected, the distinction has practical importance from an economic and from a legal perspective when it comes to protection of information (for example the protection of personal data or know-how is directed towards the semantic level; a code or algorithm could be protected or traded, either as a file or as a data stream which refers to the syntactic level) or information transfer (distinguish between 'raw' data and actual knowledge[33]; dealing with a printed book refers to the structural level). In the light of the foregoing, transferring data can have different meanings - either transferring data with a certain meaning, or transferring data as an amount of signs or transferring physical carriers containing data.

In this paper the term 'data' is used referring to the logical layer. This is because the definition provided in the DA 'reassembles data in its meaning of patterns that represent information, wherever they are reproduced – hence, data intended on the level of the logical layer'[34]. However, it must not be ignored that the three layers constitute an indissoluble whole and they

[32] Zech (n 2) 53–54; Hummel, Braun and Dabrock (n 14) 549., Y. Benkler, 'From Consumers to Users: Shifting the Deeper Structures of Regulation Toward Sustainable Commons and User Access' (2000) 52 Federal Communications Law Journal 561, 562; L. Lessig, The Future of Ideas, The Fate of the Commons in a Connected World, Random House, New York 2002, 23

[33] N. Silver, The Signal and the Noise, Penguin, New York 2012, p. 13, in Zech (n 2) 54.

[34] Cf. M. DENGA, 'Digitale Souveränität durch Datenprivatrecht?', GRUR 2022, (1113) at 1118 in Simon Geiregat, 'The Data Act: Start of a New Era for Data Ownership?' (8 September 2022) 20 <https://papers.ssrn.com/abstract=4214704> accessed 13 June 2023.

are interconnected – for example, data's true value stems from the value of the information and knowledge that can be extracted from its combination and aggregation[35], when one talks about distribution of data as a resource, he means allocation of knowledge in a broad sense, further, the access to knowledge can be hampered through adoption of technical measures. Therefore, content layer is central when it comes to valuation of data.

I.2. Types of Data

The term 'data' is not only multilayer but also a generic and therefore some further elucidations are needed.

I.2.1. Raw & Generated Data

On the basis of the categorization of data introduced by the OECD[36], some authors[37] class data according to its origin – that is, the manner in which it originated – and distinguish between raw data (provided and observed data) and generated data (derived and inferred data).

Raw data ("user-generated data") could be either provided or observed data. Provided data is data originating from the direct actions of individuals (e.g., creating user profile, registration form filing, product purchases with credit card, etc.). Observed data is data recorded by the data controller (e.g., data collected by sensors)[38].

[35] D. Rubinfeld and M. Gal, "Access Barriers to Big Data", Arizona Law Review, 2017, vol. 59, p. 342, in Tombal (n 4) 4.
[36] "Protecting Privacy in a Data-Driven Economy: Taking Stock of Current Thinking" (OECD, 21 March 2014), https://perma.cc/AFH5-MZF9 refers to provided, observed, derived, and inferred data – inferred data being defined as the "product of probability-based analytic processes".
[37] Frederike Zufall and Raphael Zingg, 'Data Portability in a Data-Driven World' in Ching-Fu Lin, Shin-yi Peng and Thomas Streinz (eds), *Artificial Intelligence and International Economic Law: Disruption, Regulation, and Reconfiguration* (Cambridge University Press 2021) 216 <https://www.cambridge.org/core/books/artificial-intelligence-and-international-economic-law/data-portability-in-a-datadriven-world/F445EC4A9E9665A05E773A88E8840027> accessed 20 May 2023.
[38] ibid.

Generated data ("data controller-generated data") encompasses derived and inferred data. Derived data is "data generated from other data, created in a 'mechanical' manner using simple, non-probabilistic reasoning and basic mathematics for pattern recognition and classification creation (e.g., customer profitability as a ratio of visits and purchases, common attributes among profitable customers)"[39]. Inferred data is "data generated from other data either by using probabilistic statistical models for testing causal explanation ('causal inferences') or by using machine learning models for predicting output values for new observations given their input values ('predictive inferences')"[40].

This distinction is important on several counts.

First, from an economic perspective, generally speaking "generated data is of higher value than raw data"[41]. Also, in relation to the valuation of data, it is to note that "on a large scale, the value of raw data increases linearly, whereas the value of generated data increases exponentially"[42] and this is a factor which must also be taken into account when providing economic analysis of a possible legal framework for data rights and data access. The dependence of the raw data's value on its amount is especially relevant when we distinguish individual-level data from aggregated data (Big Data).

Second, free flow of non-personal data is often associate with free flow of non-personal raw data only[43]. The other types of data (derived and inferred) are generated by data controllers and there is alleged economic justification for restriction of the access to them – to protect the incentives for their generation and prevent 'free-riding'[44].

[39] ibid 216.
[40] ibid.
[41] ibid 219.
[42] ibid.
[43] This is the approach of the DA proposal.
[44] However, it must be kept in mind, that in practice very often raw data and generated data are mixed and the different legal regimes could constitute a major impediment to the attainment of the free flow of data goal.

Third, as a rule the type of data reveals a stronger or weaker relationship between information and data subject – '[t]he stronger the relationship, the more individuals are involved in the creation of the data'[45], e.g., the nexus between information and the data subject can be classified as either strong (provided data), intermediate (observed and derived data), or weak (inferred data)[46]. On the one hand, a strong connection between information at stake and the data subject – natural person – may be an indication that the data constitute personal data and therefore the special rules are applicable. On the other hand, the involvement in the generation of data could be seen as a justification for granting access to or rights over the data in the context of the debate over the allocation of rights[47].

While the distinction between raw and processed data is usually considered important, it rises some interesting debates. First, some authors [48] reasonably argue that the designer or manufacturer of the IoT device can hardly be considered as sole generator of the processed data because even in smaller share the user of the device contributes to the generation of processed data. Second, they note that data collected through IoT products are often quickly analysed and processed to draw additional information from the first-level (sensor) data[49].

What is more, this initial processing may often take place within the product (using embedded software) and form direct result of a related to the use of the product service[50]. It is to be noted also that manufacturers of IoT products receive remuneration (as part of the paid price) for the embedded software, which frequently processes the raw data. These issues cast doubts on the

[45] Zufall and Zingg (n 37) 217.
[46] ibid.
[47] Martina Eckardt and Wolfgang Kerber, 'Property Rights Theory, Bundles of Rights on IoT Data, and the Data Act' (26 February 2023) 18 <https://papers.ssrn.com/abstract=4376833> accessed 21 May 2023.
[48] Josef Drexl and others, 'Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act)' (25 May 2022) 12 <https://papers.ssrn.com/abstract=4136484> accessed 11 June 2023.
[49] ibid 10.
[50] ibid.

understanding that the distinction between raw and processed data can be used unequivocally as a factor for the allocation of rights over data.

I.2.2. Co-generated Data

With regards to the issue of data generation, it must be noted that a great amount of data originates from different sources – it could be created by several machines functioning in sequence or parallel, in sequence or parallel activities of several human beings or it could be both machine and human generated, e.g., the data generated by the use of IoT. The mixture might result from both passive and active events[51]. In other words, these data are co-generated.

The notion of co-generated data is key on two grounds. On the one hand, it can reasonably be expected that in the foreseeable future co-generation will be the usual way how data are created[52]. On the other hand, on the base of the ALI-ELI principles[53] and the property law principles (related to the institutes of accession, commingling and specification in property law) some scholars[54] hold that whether and how a party should be granted an interest in the end product (co-generated data) could be determined based on an evaluation of the extent to which a party has contributed to the creation of the end product. When different parties have contributed to the creation of the end product in a very similar or equal way, in the sense that none of the contributions really outweighs the others, the researchers suggest following the example of the rules on accession and commingling that data rights should be granted to all parties involved[55].

[51] Erp and Swinnen (n 26) 69.
[52] ibid.
[53] Principle 18 defines when data are co-generated, while Principle 19 allocates the rights over co-generated data.
[54] Erp and Swinnen (n 26) 69.
[55] ibid 70.

I.2.3. Personal and Non-personal Data

Last but not least, on semantic level one important further subdivision of the generic term data is needed – data can be personal and non-personal. The EU legislation provides a detailed regulation of the personal data - GDPR[56], focused on its protection, but stipulating also rules for data access (Art. 15) and data portability (Art. 20). The division between personal and non-personal data rises very challenging questions – what does 'non-personal data' actually mean; how can non-personal data be separated from personal data in practice in cases where both types of information are generated, stored, processed etc, in conjunction in order the different legal regimes applicable to these two types of data to be enacted; to what extent is it justified to treat personal data as *res extra commercium*. In view of the limits of the current paper I will confine myself to only briefly present some of the relevant aspects of the broad discussions which the quoted problems provoke.

First, it is clear that a dichotomy between personal and non-personal data originates from the fact that with the GDPR within the EU is established a comprehensive protection regime applicable exclusively to personal data. However, legal researchers across Europe in unison identify as a challenge the existence of much legal uncertainty about where to draw the line between personal and non-personal data (and how to deal with mixed data sets)[57]. The difficulty of the task lies in the fact that the definition of personal data is broad, making it a context-specific, dynamic, fluid and open-ended concept[58]. In this context the approach adopted be the EU legislator not to define 'non-personal data' but to rely on a negative, residual definition

[56] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 4.5.2016

[57] Eckardt and Kerber (n 47) 24.

[58] Somaini (n 13) 89; Thomas Tombal and Inge Graef, 'The Regulation of Access to Personal and Non-Personal Data in the EU: From Bits and Pieces to a System?' (15 December 2022) 4–5 <https://papers.ssrn.com/abstract=4304148> accessed 21 May 2023.

based on the term 'personal data'[59] does not provide legal certainty. On the contrary, it attempts to explain one unknown term with another broad and vague concept. Second, it is known that at least theoretically one could circumvent the application of the GDPR regime by anonymising the personal data before sharing them[60]. However, even if this might be possible in some cases, it is believed that there are other cases where this might reduce the value of the dataset[61]. Furthermore, some experts argue that truly effective anonymisation is difficult to be achieved, especially in light of the constant development of Big Data analytics, which increase the risk of reidentification of the data subjects. On the basis of the demonstrated failure to effectively anonymise personal data, these experts conclude that 'what is often presented as anonymisation techniques are, in fact, merely pseudonymisation techniques. Yet, pseudonymised data remain personal data covered by the GDPR, given that the data subject can still be re-identified'[62].

Having in mind the underlined limitations of the anonymization process, some scholars conclude that the traditional distinction between personal and non-personal data has a rather 'porous boundary' and thus propose a new data typology, based on the way in which the data has been generated or acquired, rather than on the classic distinction between personal and non-personal data[63]. It should be noted also that the classic distinction is based on the understanding that personal data and non-personal data are completely different phenomena and therefore they deserve completely different regulation. However, the belief that personal data is no one's property is illusional[64] - actually there is enough grounds to accept that 'the Information

[59] See Art. 2(1) Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union, PE/53/2018/REV/1, OJ L 303, 28.11.2018

[60] Tombal (n 4) 17; 'Shall We Share? The Principle of FRAND in B2B Data Sharing by Marco Botta :: SSRN' 12 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4423805> accessed 21 May 2023.

[61] Tombal (n 4) 17.

[62] ibid.

[63] Thomas Tombal, "Imposing Data Sharing among Private Actors." (Kluwer Law International, 2022). Chapter 1.2 – 1.3, in 'Shall We Share? The Principle of FRAND in B2B Data Sharing by Marco Botta :: SSRN' (n 60) 12–13.

[64] Nadezhda Purtova, 'Illusion of Personal Data as No One's Property' (29 October 2013) <https://papers.ssrn.com/abstract=2346693> accessed 24 May 2023.

Industry actors presently have de facto property rights in personal data'[65]. In this context, the sensible approach is to acknowledge the differences between these two types of data and to take them into consideration when developing the applicable legal regime but not to exclude the personal data from the debates over the rights over data and the data access.

Despite the fact that the criticisms seem to be substantiated and the suggestions appear sensible and therefore they deserve support, we have to apply the legal rules as established. Which, however, leads to the next issue - the emerging EU legal framework on data sharing, specifically the proposed DA, has a wide scope of application covering both personal and non-personal data and this may in part be because in some cases it is practically impossible or excessively difficult to separate the two types of data. However, the proposed DA insist on the compliance with the protection of personal data rules under the GDPR[66], e.g., where in a data set personal and non-personal data are inextricably linked the data protection rules must apply to the whole set.

In this context several authors[67] acknowledge that, on the one hand, personal data protection considerations are used by large data holders to justify refusals to share data with third parties[68], and on the other hand, these large data holders can better absorb the large implementation costs

[65] ibid 28.
[66] See recitals 7 and 30 and Art. 1 (3) DA
[67] See J. Crémer, Y.-A. de Montjoye and H. Schweitzer, "Competition Policy for the digital era", op. cit., p. 99. See also M. Gal and O. Aviv, "The Competitive Effects of the GDPR", Journal of Competition Law and Economics, September 2020, Volume 16, Issue 3, p. 349-391; D. Geradin, T. Karanikioti and D. Katsifis, "GDPR Myopia: How a Well-Intended Regulation ended up Favoring Google in Ad Tech", TILEC Discussion Paper DP 2020-012, May 2020, available at https://ssrn.com/abstract=3598130; J. Campbell, A. Goldfarb and C. Tucker, "Privacy Regulation and Market Structure", Journal of Economics & Management Strategy, vol. 24, issue 1, 2015, p. 47-73; J. Jia, G. Zhe Jin and L. Wagman, "The Short-Run Effects of GDPR on Technology Venture Investment", November 2019, available at https://papers.ssrn.com/abstract=32789128; T. Zarsky, "Incompatible: The GDPR in the Age of Big Data", Seton Hall Law Review, 2017, Vol. 47, No. 4(2), p. 995- 1020; T. Zarsky, "The Privacy–Innovation Conundrum", Lewis & Clark Law Review, 2015, Vol. 19, No. 1, p. 115-168. For empirical evidence of this increased concentration, see G. Johnson and S. Shriver, "Privacy & market concentration: Intended & unintended consequences of the GDPR", March 2020, available at https://ssrn.com/abstract=3477686 – in Tombal (n 4) 19.
[68] Applying at the same time 'double standards', i.e. the large data holders adopt a very restrictive approach towards data sharing with third parties while massively circulating their users' data internally - ibid 19–20.

of the GDPR than smaller competitors, which could ultimately create some serious competition issues. Therefore, legal scholars conclude that the GDPR reduces competition, increases concentration, and potentially strengthens big data holders in data and data-related businesses. It also enhances the EU's already existing hurdles to data sharing, thereby decreasing data synergies that could result from integrating disparate datasets managed by distinct bodies[69].

In other words, the personal data protection regulation not only constitute a legal or regulatory barrier before the free flow of data, but it is also used as an excuse for the adoption of anti-competitive measures from large market players.

To summarize, the term ‘data’ has several layers. The current paper refers to the logical level of information (without ignoring the interrelation between the three layers). On this level ‘data’ constitutes a generic term and different sub-types of information can be distinguished. These sub-types are usually connected, but they have different economic importance and legal regimes, which shape the features of the global digital ecosystem.

I.3. Data as a Resource

To understand whether we should regulate the access to data and the rights over data and in order to formulate what exactly types of access and rights are needed, we should explore the data as a resource and assess its role for the economy. One should bear in mind, however, that data is a complex good, and this is actually one of the factors which makes determining the legal framework to be applied to data a complex task[70].

[69] M. Gal and O. Aviv, “The Competitive Effects of the GDPR”, op. cit., p. 352, in ibid 19.
[70] ibid 1.

It is known that one of the main critics to the DA proposal is for the lack of clear economic analysis[71]. The current contribution without claiming to provide in-depth economic analysis tries to take into account the interrelation between law and economics.

Before all else, when we analyze data as resource, we should bear in mind that 'data have no value on their own. They only become valuable to the extent that parties can use them to leverage their position in data-driven services markets'[72], which means that the importance of the data will depend to great extent on the market and on the level of economic and social development of the society. This in general means that with the development of Industry 4.0 the value of data can be expected to increase.

I.3.1. (Non-)rival and (Non-)excludable resource

One of the most important characteristics of data is that it is non-rival resource, which means that same data can be utilized by unlimited group of subjects simultaneously. It is known also that 'the same data could also be shared and collected by different entities without depleting the source of data for others'[73]. At the same time, some experts identify that 'observed user data collection (as opposed to volunteered user data) is rival […] because for key services […] the market is concentrated with only a few firms able to track user activity across the web. Thus, observed data is not ubiquitously available, and it is also usually neither feasible nor socially desirable to duplicate the collection of the same observed data'[74], i.e., it is to be noted

[71] Eckardt and Kerber (n 47) 3.
[72] Bertin Martens and others, *Business-to-Business Data Sharing: An Economic and Legal Analysis* (European Commission 2020) 14.
[73] Alexandre Streel, Jan Krämer and Pierre Senellart, 'Making Data Portability More Effective for the Digital Economy' [2021] SSRN Electronic Journal 51.
[74] ibid 8.

that even though the acts, facts or information which generate or constitute data could have no limits in range or scope, the collection of observed data is de facto limited and thus often rival.

In addition, analyzing the characteristics of personal data as a resource some scholars conclude that its rivalrous nature follows from the rivalrous nature of the users of electronic platforms and the time and effort, they spend on digital platforms[75].

These arguments may be transmitted to non-personal observed data too - the rivalrous nature of the users of IoT devices and the time and effort they spend operating with them presumes the rivalrous nature of the generated data. However, this is actually sui generis rivalrous nature of a resource – the rivalry does not occur when using the product (unlimited number of users can utilize it simultaneously), but when producing it (generating data).

This peculiarity allows parallel with the creation and the use of IP objects, e.g., while an invention can be used simultaneously as non-rivalrous resource, the creation of such invention is rivalrous. Here, the objective of the patent protection is to preserve the rivalrous nature of the creation process, to incentive generation of knowledge, and the IP law achieve this by the imposition of restrictions on the use of the invention which theoretically can be used otherwise as a non-rivalrous resource. The question, therefore, is whether with regards to data there is a justification for restricting the use of the product only because the process of its creation is rivalrous. The answer depends on whether without such a protection the generation of data would possibly be reduced.

The non-rival nature of data allows the data holder to grant either 'access' or 're-use' to a third party, without losing the right to the continuing use of the same dataset[76]. It is argued that while

[75] Purtova (n 64) 25.
[76] 'Shall We Share? The Principle of FRAND in B2B Data Sharing by Marco Botta :: SSRN' (n 60) 21.

data is in principle a 'non-rival' resource[77], it is de facto excludable - not by nature, but in practice through both technical and contractual instruments[78]. What is more, as pointed out by some scholars "data's competitive significance (and value) arises in part from the ability of firms to exclude others from access and analyzing it as quickly"[79]. It is to be noted that the ICTs dramatically reduce the cost of cost of exclusion of digital data[80]. Only when a good is both non-rivalrous and non-excludable, it can be considered to be a public good.

I.3.2. Non-depletable Resource with Lossless Acquisition and Use

Data are also non-depletable: they can be used more than once without losses in quality[81]. Some authors claim that in some exceptional cases information can be used only a single time[82], however, from the exemplification of this hypothesis, which they provide - information where oil is located under a particular parcel of land, is obvious that they are not actually the data that can be used only once (data can be used numerous times, e.g., for cartographical, analytical and other purposes), but the resource to which the data at stake point – oil resources. Furthermore, acquisition and usage of information are lossless - 'contrary to other things that

[77] Some scholars argue that data are not always and, in every circumstance, non-rivalrous. They state that some data can lose their value as soon as e.g., illegitimate users have access to them, at least for those that wish to protect the information. In particular cases the consumption of data would affect their potential to meet the demands of (a few) others - Christian Reimsbach-Kounatze, 'Enhancing Access to and Sharing of Data: Striking the Balance between Openness and Control over Data' (2021) 33. Others explain that 'because information helps construct communities, those who possess information possess a form of power. Sharing information diminishes the power held by the person who previously restricted access to the information. Information therefore is socially rivalrous' - David W Opderbeck, 'Socially Rivalrous Information: Of Candles, Code, and Virtue' (21 August 2007) 85 <https://papers.ssrn.com/abstract=1008500> accessed 27 May 2023.
[78] Tombal (n 4) 6.
[79] M. Stucke and A. Grunes, Big Data and Competition Policy, op. cit., p. 46, in ibid.
[80] Reimsbach-Kounatze (n 77) 33.
[81] Hummel, Braun and Dabrock (n 14) 549.
[82] Louis Kaplow and Steven Shavell, 'Economic Analysis of Law' (1 February 1999) 26 <https://papers.ssrn.com/abstract=150860> accessed 6 June 2023.

one owns, one's personal information is not lost when acquired by someone else'[83]. In addition, reproduction of data is cheap and the sharing – easy[84].

I.3.3. Economies of Scale and Scope, Indirect Network Effect

Further, data could be characterised by important dynamic economies of scale, which means that large firms with large amounts of data have the opportunity to raise product quality at lower costs than small firms, and economies of scope – big data holders can leverage the data, or the insights due to machine learning, that they receive from an existing service or good to enter into an adjacent market with a higher quality product, demonstrating a novel form of economies of scope[85]. Economies of scope have one additional manifestation: economies of scope in data aggregation, e.g., 'when two datasets are complementary, more insights and economic value can be extracted from merging them, compared to keeping them in separate data silos'[86], which could constitute an incentive for cooperation between data holders. These economies of scale and scope are believed to create 'data-driven indirect network effects'[8788]: a difference in the amount and quality of the obtained by an actor on data-driven market data tend to tip the market in favour of this actor, and when such a market tips, it is difficult to re-establish competition other than granting access to data.

[83]Floridi, L. (2014a). The fourth revolution: how the infosphere is reshaping human reality. Oxford: Oxford University Press, 118, in Hummel, Braun and Dabrock (n 14) 558.
[84] Maria Jose Schmidt-Kessen, 'The Impact of Data Ownership Rights on Competition in Big Data Markets: Reflections in the Context of the EU and Global Data Economy' (2020) 2 VIT Law Review 56, 69.
[85] Scott Morton, F., Bouvier, P., Ezrachi, A., Jullien, B., Katz, R., Kimmelman, G., Melamed, D. & Morgenstern, J. (2019). Committee for the Study of Digital Platforms: Market Structure and Antitrust Subcommittee Report. Draft. Chicago: Stigler Center for the Study of the Economy and the State, University of Chicago Booth School of Business, 14.
[86] Martens and others (n 72) 12.
[87] Some scholars claim that the manufacturers of IoT do not benefit from network effects - Drexl and others (n 48) 16.
[88]'Competing with Big Data* - Prüfer - 2021 - The Journal of Industrial Economics - Wiley Online Library' 969 <https://onlinelibrary.wiley.com/doi/full/10.1111/joie.12259> accessed 27 May 2023.

It is also believed that data-driven indirect network effects can give rise to self-reinforcing feedback effects from a dynamic perspective[89]. One of the main feedback loops consists in the fact that the more consumers are using a service or a good, the more (volunteered and observed) data is created on which data analytics can be performed and algorithms can be trained, which in turn results in an improvement of the good or the service, which in turn leads to more consumers (user feedback loop)[90]. The larger number of users normally leads to an increase in the firm's revenues.

I.3.4. Infrastructural Resource

From an economic perspective data may be considered as an 'infrastructural resource'[91] because data are means rather than ends, and their demand is driven by the demand for the derived outputs. Data are thus factors of production[92]. It is recognized that assets can be classified as infrastructures when they meet the following criteria: '(1) The resource may be consumed non-rivalrously for some appreciable range of demand. (2) Social demand for the resource is driven primarily by downstream productive activities that require the resource as an input. (3) The resource may be used as an input into a wide range of goods and services, which may include private goods, public goods, and social goods' [93].

I.3.5. Key Capital Good

Data are now seen as 'key components ('inputs') for new innovative technologies, capable of improving products and processes and to support decision-making of public bodies as well as

[89] 'The Role of Data for Digital Markets Contestability: Case Studies and Data Access Remedies' (*CERRE*) 64 <https://cerre.eu/publications/data-digital-markets-contestability-case-studies-and-data-access-remedies/> accessed 27 May 2023.
[90] 'The Role of Data for Digital Markets Contestability: Case Studies and Data Access Remedies' (n 89).
[91] Reimsbach-Kounatze (n 77) 31.
[92] ibid 34.
[93] Frischmann, B. M. 2012. Infrastructure: The Social Value of Shared Resources: Oxford University Press, 61, in Fia (n 6) 201.

of private actors'[94] and thus the access to data is 'a key factor to the emergence of a data economy'[95]. More precisely, data are capital good because data are not used up, exhausted, or otherwise transformed when used[96]. In the context of the European attempts to unlock the potential of the digital economy through implementation of mechanisms for access, sharing and re-use of data[97], it is crucial to mark that data are general-purpose input because data can be used according to capabilities of the users for a wide range of private, public and social goods and services[98].

***

To briefly summarize, data as a resource is infrastructural, non-rivalrous (but de facto excludable), non-depletable resource, which acquisition and usage are lossless. Data is also characterized by dynamic economies of scale and economies of scope, which create data-driven indirect network effects.

II. Distribution of Data Wealth

In this chapter the political economy of non-personal data which currently exists is present. Further the chapter continues with a brief exposition of the existing concepts for allocation of the benefits of non-personal data. For sake of brevity, I will focus only on the two main conceptual extremes - free and unrestricted access to data and full enclosure of data. This is needed in order to contextualize both – the status quo, which is also shortly explained, and the

[94] Francesco Mezzanotte, 'Access to Data: The Role of Consent and the Licensing Scheme' [2017] S. Lohsse, R. Schulze, D. Staudenmayer (eds.), Trading Data in the Digital Economy: Legal Concepts and Tools, Hart/Nomos, Oxford/Baden-Baden 159 <https://www.academia.edu/37037356/Access_to_Data_The_Role_of_Consent_and_the_Licensing_Scheme> accessed 28 May 2023.
[95] ibid 160.
[96] Reimsbach-Kounatze (n 77) 34.
[97] See EC, A Digital Single Market Strategy for Europe, COM(2015) 192 final, 6 May 2015; cf. also EC, A European strategy for data, COM(2020) 66 final, 19 February 2020.
[98] Reimsbach-Kounatze (n 77) 35–36.

political economy model constructed under the DA, which assessment is the topic of the next chapter.

II.1. Data Access

Having acknowledged what is the meaning of 'data' and what are the economic characteristics of data as a resource the research can proceed with the problem how the wealth of this resource can be distributed and what legal measures can be adopted in order the fair allocation to be guaranteed.

The analysis of this problem implies a need to clarify the concept of data access. On the first place, it must be clear that access takes place on different degrees depending on how access opportunities are distributed[99].

Illustratively data access can be pictured as a vertical line or pyramid whose summit is open access i.e., providing indiscriminate access to the public, and whose other end is closed access under which seekers cannot access data[100].

For the purposes of this work, the political economy of data in the two conceptual extremes (free and unrestricted access and full enclosure), as well as in the current pre-DA reality, shall be examined in order. This is needed because any data access legal scheme follows an approach, which is the legal and technological construct of the rules revealing access seekers' prerogatives[101]. It is to be noted also that 'this criterion mingles a pure legal and coercive dimension, based on duties and rights, and a technological one, through which rules are embedded into the technological architecture of data processing systems'[102]. In order for an

[99] Fia (n 6) 189.
[100] ibid.
[101] ibid 190.
[102] Lessig, L. 2006. Code: Version 2.0: Basic Books, in ibid.

effective system for fair distribution of data, which would in all likelihood will position access opportunities somewhere in the grey area between the two extremes, to be constructed, it is needed the strengths and weaknesses of the conceptual extremes and the present situation to be acknowledged and assessed.

II.2. Data as a Digital Public Good

In view of the economic characteristics of data as a resource and the fact that 'data has become the lifeblood of our economy'[103] it seems natural the access to non-personal data to be 'free as the air we breathe'[104]. In fact, the economics-of-information literature often uses information as a perfect example of a public good[105]. With regards to the characteristics of data it is even claimed that 'the benefits of data sharing may arguably be greater than the benefits of sharing other resources, and the costs of data sharing may arguably be smaller than the costs of sharing other resources'[106].

Numerous arguments that can be provide in support of a public domain regulation[107] of machine-generated datasets have been acknowledged[108]. Scholars have formulated two main reasons substantiate why commons management of raw non-personal data may be desirable – on the one hand, these data can be deemed a cooperative infrastructural resource that calls for being pulled out of its factual enclosure to open up the benefits of data-driven innovations to a greater number of actors ('structuralist approach'), on the other hand, grasping these data as a

[103] Jane Yakowitz, 'TRAGEDY OF THE DATA COMMONS' 25 42.
[104] 'The Right to Process Data for Machine Learning Purposes in the EU by Mauritz Kop :: SSRN' 13 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3653537> accessed 30 May 2023.
[105] Purtova (n 64) 19.
[106] Tombal (n 4) 16.
[107] Some authors refer to commons management of data as 'data socialism' (Kyung Park, 'Data as Public Goods or Private Properties?: A Way Out of Conflict Between Data Protection and Free Speech' (2021) 6 UC Irvine Journal of International, Transnational, and Comparative Law 77, 99.), however, this may give the erroneous impression that such approach could detriment market competition while on the contrary – access to data is a factor for enhancement of competition and innovation.
[108] Mezzanotte (n 94) 171., see also 2015. Data-Driven Innovation. Big Data for Growth and Well-Being. Paris: OECD Publishing, https://doi.org/10.1787/9789264229358-en

commons means valuing their functional nature, making data available to a wide number of actors for the fulfilment of fundamental rights and enhancing human flourishing ('functionalist approach')[109].

For the sake of brevity, I will only confirm the understanding that 'the open data approach can work as a default rule governing RNPD'[110] since wider availability of data has the potential to ensure 'contestability, fairness and innovation and the possibility of market entry, as well as public interests that go beyond competition or economic considerations'[111].

It is important, however, to note some concerns, related to this approach.

First, the argument that raw non-personal data are, on the one hand, result of cooperative processes[112], and on the other – just a by-product or a side effect of activities in which many actors engage[113], is relevant only when it comes to observed data (or at best to data processed through embedded in IoT software). However, as it was noted above, generated data are much more valuable than observed data. Therefore, an approach focused exclusively on raw data may be too narrow and unable to ensure unleash of the economic potential of data.

Furthermore, and here we should turn to the code layer of the notion of data, one of the main obstacles to use of data is the lack of standardization and interoperability of tools and format of data storing and transfer. Therefore, the open access to data without incentive for the data holders to ensure standardization and interoperability would not be sufficient to satisfy the needs of data users.

[109] Fia (n 6) 200.
[110] ibid 206.
[111] Communication from the Commission, "Shaping Europe's digital future", op. cit., p. 9; Communication from the Commission, "A European strategy for data", op. cit., p. 5 and 14
[112] Fia (n 6) 200.
[113] ibid.

These two issues indicate that the legislator should adopt the carrot and stick approach with the data holders and attempt to enhance cooperation on the data market, providing both sweeteners for coopetition and measures to overcome unjustified data enclosures.

Finally, it must be clear that data are infrastructural resources with 'dual-use' nature[114], which can be utilized by nefarious actors. For that reason, the system which distributes the access to data has to provide also traceability and accountability in order to prevent malicious use of the resource. To achieve this the legislation could entrust a neutral data trustee that grants access to and shares the IoT data[115]. However, it may be more feasible to achieve traceability and accountability through assignment of duties to data holders in exchange for some benefits.

II.3. Exclusive Data Ownership

The term 'data ownership' has acquired popularity not only among laypeople but in the legal literature too[116]. This allows analogy with classic property rights, despite the fact that the term itself is not precise, because rights over data are not included in the *numerous clausus* of property (including IP) rights and therefore there can be no actual such rights over data[117], and additionally some confusion is possible, since the title 'ownership' has established specific meaning which does not fully mirror the concept of exclusive control over data, interest in an object with erga omnes effect. What is initially framed as data ownership concerns primarily controllability i.e., the availability of effective means for data subjects to exercise control over data at stake[118].

[114] Tommaso Soave, 'Digital Humanitarians and International Lawyers: Worlds Apart or Two Sides of the Same Coin?' (2022) 25 Max Planck Yearbook of United Nations Law Online 718, 740.
[115] The idea for data trustees is well-known – see Eckardt and Kerber (n 47); Kerber (n 8); ibid.
[116] See Fia (n 6); Somaini (n 13); Hummel, Braun and Dabrock (n 14).
[117] 'The Right to Process Data for Machine Learning Purposes in the EU by Mauritz Kop :: SSRN' (n 104) 6.
[118] Hummel, Braun and Dabrock (n 14) 554.

It is known that the property rights theory deconstructs 'property' into a bundle of rights over a resource[119] i.e., a bundle of possible uses assigned to the rights holder[120]. It can be argued that with some modifications the three pillars of classic *jus in rem*:0 *usus* (right to use, to access), *fructus* (right to profit) and *abusus* (right to dispose and to deny use and proceeds to the others) can serve as a basis for the construction of rights over data[121].

Undoubtedly these rights are developed to fit the needs for regulation of the rights over classical corporeal objects which unlike data are normally rivalrous goods with exclusive use. However, nowadays it is established that data can be 'owned'[122]. The differences between tangible and intangible objects do not significantly impact the legal and economic status of the owner and her opportunities – generally she has the right to benefits economically from the object and to exclude the others from it.

The IP rights over intangible assets can be deconstructed to similar to in rem rights e.g., patents limit the use of information without limiting access and copyright limits the information by limiting access[123] i.e., patent holders are granted with fructus and quasi-abusus (they cannot destroy the object of protection but can deny use and proceeds to the others) and copyright holders generally have quasi-usus, -fructus and abusus (they can restrict the access to the object, they can also profit from it and deny use and proceeds to the others without the opportunity to destroy the knowledge). Similarly, the Database Directive[124] provides copyrights and sui generis rights for the creators of databases.

[119] Eckardt and Kerber (n 47) 4.
[120] Zech (n 2) 56.
[121] See ibid 55–56.
[122] Park (n 107) 94.
[123] Zech (n 2) 57.
[124] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases

Once it is clear that (quasi-) exclusive rights over resources with similar to data characteristics can be constructed we may proceed with the question for the justification of property. Scholars distinguishes three questions - general: why should there be property? specific: what kind(s) of property rights should there be? particular: who should have a title to a specific kind of property?[125]

Although some scholars advocate for creation of rights over data for various reasons[126], the voices against introduction of data ownership dominates[127] in the literature mainly because this 'would increase transaction costs and impede the trading and the use of data'[128]. The Commission has also abandoned its initial idea for 'data producer's right'[129] and now advocates for access rights and similar data rights[130]. It is established that the burdens of introducing and enforcing additional legal mechanisms to govern resources should be proportional to the societal benefits these mechanisms generate. This rationale is discussed for data ownership as well[131].

In light of the above we should briefly answer the question why should there be property over data, would it generate significant societal benefits? This analysis requires to assess whether

[125] Becker, L. C. (1980). The moral basis of property rights. Nomos XXII: Property, 22, 187–220, in Hummel, Braun and Dabrock (n 14) 551.
[126] See Kish, L. J., & Topol, E. J. (2015). Unpatients—why patients should own their medical data. Nature Biotechnology, 33, 921, 923 in ibid 556., Thouvenin, F. (2017). Wem gehören meine Daten? Zu Sinn und Nutzen einer Erweiterung des Eigentumsbegriffs. Schweizerische Juristen-Zeitung, 113, 21–32., in ibid 446., M. Amstutz, 'Dateneigentum: Funktion und Form', Archiv für die Civilistische Praxis 218 (2018), 439–551, at 489 et seqq.; see also F. Cheneval, 'Property Rights of Personal Data and the Financing of Pensions', Critical Review of International Social and Political Philosophy (2018), 1–23; I. Landreau et al., 'My Data Are Mine: Why We Should Have Ownership Rights on Our Data' (Paris: GenerationLibre, 2018), at 18 et seqq.; E. Tjong Tjin Tai, 'Data Ownership and Consumer Protection', Journal of European Consumer and Market Law 7 (2018), 136–140, at 136 et seqq., in Thouvenin and Tamò-Larrieux (n 21) 321. Zech (n 2) 74.
[127] Park (n 107) 99; Daniel Zimmer, 'Property Rights Regarding Data?' (2017) 106.
[128] Thouvenin and Tamò-Larrieux (n 21) 338.
[129] Communication on 'Building a 46 European Data Economy' (COM(2017) 9 final, p. 10 ff)
[130] COM(2020) 66 final 3 p. 4 ff.; COM(2018) 232 final, p. 9
[131] Hummel, Braun and Dabrock (n 14) 549. See also the quoted there sources.

data ownership is necessary to address market failures, and in general in what (if any) respects does the protection of data as object of exclusive property promote social welfare?

When it comes to market failures which impose protection based on property rights one can distinguish a market failure in a narrow sense from a market failure in a wider sense[132]. In a narrow sense, a market failure arises if the good would not be produced or used unless there were property titles in it, and in a wider sense, a market failure arises if the transaction costs are not as low and allocations not as efficient as they could be[133].

Market failures in narrow sense are traditionally acknowledged as justifications for the protection of property – protection of the rights over a good serves as an incentive for its production[134]. However, currently no incentive problem can be identified with regard to the generation of non-personal raw data[135]. On the contrary, it is known that nowadays the amount of data generated is far greater than it ever has been - in the Industry 4.0 codified information is generally produced and collected at very low costs, often as by-products of core business activities[136]. It speaks volumes that by 2025, the amount of data generated globally will reach 175 zettabytes[137]. With regards to the above the assumption that ownership rights would create incentives for data-related products and services was deemed 'highly speculative'[138].

In my opinion, market failure in narrow sense could be generally identified also in the hypothesis in which without establishment of rights over unowned objects there is a risk of

[132] Florent Thouvenin, 'Wem gehören meine Daten? Zu Sinn und Nutzen einer Erweiterung des Eigentumsbegriffs' (2017) 113 Schweizerische Juristen-Zeitung 21. In Hummel, Braun and Dabrock (n 14) 549–550.
[133] Thouvenin (n 132). In Hummel, Braun and Dabrock (n 14) 549–550.
[134] Kaplow and Shavell (n 82) 14.
[135] Kerber (n 8) 128.
[136][136] Kerber (n 36) 993, reporting OECD, 'Data driven innovation' (2015) 185. Available online under: accessed 14 June 2017); but see already Ayres, Super Crunchers. Why Thinking-by-Numbers Is the New Way to Be Smart (Sperling & Kupfer 2008) 60, as quoted in Mezzanotte (n 94) 172.
[137] Fia (n 6) 182.
[138] Max Planck Institute for Innovation and Competition, 'Arguments Against Data Ownership. Ten questions and answers', https://www.ip.mpg.de/fileadmin/ipmpg/content/forschung/Argumentarium-Dateneigentum_eng.pdf

exhaustion of certain resource e.g., establishment of rights over hunting fields to prevent depletion of animals, rights over ocean's fisheries and the sea bed to preserve the resources of the sea (fish, oil and mineral resources), or rivalrous resources e.g., the electromagnetic spectrum[139]. However, from the characteristic of data as a resource (non-rivalrous, non-depletable, lossless acquisition and usage) it is obvious that the quoted concerns, relevant for other resources, cannot serve as justifications for establishment of exclusive rights over data.

When it comes to market failure in the wider sense i.e., deficiencies in the effective management of the resource, researchers argue that there is no sufficient empirical and conceptual evidence about the superiority of the ownership paradigm[140].

While no unquestionable justification for granting exclusive control over non-personal IoT data can be identified ownership rights in data rises a lot of concerns. In the literature nowadays it is commonly accepted that such a regime 'would have the potential of suffocating the European data economy rather than boosting it'[141]. Furthermore, since the consumers 'would readily contract away their ownership, very much as they are currently contracting away any other rights they have with regard to data, this is not likely to enhance consumer rights either'[142]. Such an approach is expected to ultimately lead to under-utilization of the non-personal IoT data[143].

Paradoxically, data ownership, introduced to overcome a market failure, may itself cause another market failure: the 'tragedy of the data commons'[144] The traditional tragedy of the

[139] Kaplow and Shavell (n 82) 15–16.
[140] Thouvenin (n 132). In Hummel, Braun and Dabrock (n 14) 550.
[141] Maartje Elshout et al., Study for the EC on consumer's attitudes towards terms and conditions, 2016, p. 9; Jonathan A. Obar and Anne Oeldorf-Hrisch, 18 The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of 19 Social Networking Services, (2018) 22 iCS1 in ALI-ELI Principles, Reporter's note to Principle 29.
[142] Ibid.
[143] Eckardt and Kerber (n 47) 13.
[144] Yakowitz (n 103) 4.

commons describes self-interested actors who convert the communal benefits of the commons into private benefits for themselves[145]. The tragedy of the data commons describes self-interested data subject who depletes the commons by removing his data. The marginal detriment of his decision is externalized and shared across the entire population. The results in the two hypotheses are identical: communal benefits are lost due to actions motivated by self-interest[146].

Having regard to the foregoing, it is obvious that the general question why should there be property rights over data could hardly find well-founded response. Even if some justification is recognized the following two questions are not easy too.

The hypothesis of co-generated data is particularly challenging, because it leaves open the question of the concrete design of this bundle of rights. Should there be joint ownership on this IoT data, or can the cogenerating actors use, share, and monetize the non-personal IoT data independently from each other?[147]

The arbitrary and incomprehensible allocation of exclusive rights over data has the potential to lead to situation known as 'tragedy of the anticommons' i.e., 'scenario in which too much private property results in the under-use of resources. Such is the case especially where ownership is fragmentary and absolute. In this setting, too many right-holders failing to put their resources to use, or even one right-holder neglecting to employ the resource, can result in the resource not being used at all'[148].

[145] ibid.
[146] ibid.
[147] Eckardt and Kerber (n 47) 18.
[148] David Lametti, 'The Concept of the Anticommons: Useful, or Ubiquitous and Unnecessary?' (2 January 2013) 233 <https://papers.ssrn.com/abstract=2195549> accessed 12 June 2023.

When it comes to allocation of rights over data at least two hypotheses of 'tragedy of data anticommons' are possible. First, when each data holder has exclusive rights over the data under her control the scenario similar to the so-called 'patent thicket'[149] can happen – each right holder has the opportunity to extract the value of her data only and the great potential of the data industry remains unrealised[150]. Second, if exclusive rights are granted to person who has no incentive to utilize this resource, its potential will again be locked.

To conclude, the outlined pitfalls demonstrate that when it comes to non-personal raw data the property model is 'the wrong choice, not only for efficiency reasons, but also because it fails to meet the distributional goals required for justice'[151].

II.4. The state of affairs before DA

A broad consensus exists that so far, no legal rights with respect to non-personal raw data are granted[152].

It is to be noted, however, that law and politics structure economic activity both in their absence and in their presence[153], i.e., the absence of legal rules does not mean that goods for which a market demand exists will remain undistributed. On the contrary, 'self-regulation' is 'simply name for particular kinds of regulatory and disciplinary regime[154].

[149] 'The thicket occurs where there are many patent- holders – 'owners' – who hold exclusive patent rights that are useful as part of a larger process or invention […]. In such complex clusters where a group of patents might be combined for a larger functional purpose, one owner of a key patent might block the whole invention from ever happening' - Michael A. Heller and Rebecca S. Eisenberg, 'Can Patents Deter Innovation? The Anti-Commons in Biomedical Research' (1998) 280 Science 698, in ibid.
[150] It should be reminded that on a large scale the value of data increases.
[151] Yakowitz (n 103) 63.
[152] Eckardt and Kerber (n 47) 6.
[153] Michael A Wilkinson and Hjalte Lokdam, 'Law and Political Economy' (20 March 2018) 6 <https://papers.ssrn.com/abstract=3144723> accessed 12 June 2023.
[154] ibid.

Therefore, the question is not weather but how wealth of data is allocated among the society[155]. The answer is well-known: 'currently, data collectors by default have exclusive economic rights to all the data that they collect, as long as some privacy protections are taken care of'[156]. The logic of such unilateral appropriation is, however, difficult to comprehend[157].

The de facto control in combination with the lack of legal remedies for the access seekers results in so-called 'data sharing paradox'[158]: on the one hand, data is growing in value within the data economy; access to data becomes 'the fuel of innovation and knowledge creation in an increasingly connected world'[159]. On the other hand, due to several obstacles, firms are reluctant to share the collected data; they prefer to 'seal' the data, rather than concluding a sharing agreement with a third party[160]. The data sharing paradox affects not only B2B transactions but also G2B and B2G data sharing[161].

This distribution of wealth and resources with regards to data could be explained through Umbeck's theory of formation and initial distribution of property rights, according to which in absence of property rights in a valuable scarce resource or when these rights are ill-defined[162], they will be allocated in a way that is proportionate to the ability to exclude others from that resource[163].

[155] A separate issue is the question to what extent de-facto position which is not in any way normatively justified is legitimate from economic and policy perspective.
[156] Singh and Vipra (n 1) 56.
[157] ibid.
[158] 'Shall We Share? The Principle of FRAND in B2B Data Sharing by Marco Botta :: SSRN' (n 60) 8.
[159] Tombal (n 4) 1.
[160] Alberto Alemanno (2018), 'Big Data for Good: Unlocking Privately Held Data to the Benefit of the Many.' 9 European Journal of Risk Regulation: 185., in 'Shall We Share? The Principle of FRAND in B2B Data Sharing by Marco Botta :: SSRN' (n 60) 8.
[161] ibid.
[162] For the limited types of data for which IP protection exists and for which ownership is excluded or would be severely limited see Schmidt-Kessen (n 84) 73–75.
[163] John Umbeck, A Theory of Property Rights: With Application to the California Gold Rush (Iowa State University Press, 1981) in Purtova (n 64) 5–6.

As it was clarified data is characterised by economies of scope and scale, which provide an advantage to data holders and constitutes an incentive to collect and produce as much data as possible[164]. At the same time without any duty or incentive to share innovation, data holders keep it all for themselves[165].

In practice this means that even without 'de jure' rights, the manufacturers of IoT products, can implement a technical design for their IoT devices that gives them exclusive 'de facto' control over all data generated by the use of the device by the firms or consumers who have bought, leased or rented it[166].

The de facto entitlement of a data holder to actually consume the 'good' in question equals economic property right which even without the recognition of the law, may also be self-enforced[167]. Against this backdrop, it is to be noted that in contrast to the legal property rights created by the legislative process, economic property rights represent one's de facto ability to enjoy a resource and exclude others from that resource as a result of a complex interaction of factors such as the effectiveness of law enforcement and other non-proprietary legal arrangements[168].

With regard to the above, it is established that de facto data ownership brings about at least two failing[169]s. On the one hand, due to imbalances in negotiating power[170] neither the users of IoT products, nor the other firms can get enough access to this data[171]. On the other hand, the de facto enclosure of needed data is widely recognised as a hindrance to innovation in Europe.

[164] B. Martens, A. de Streel, I. Graef, T. Tombal and N. Duch-Brown, "Business to business data sharing", op. cit., p. 13. In Tombal (n 4) 6.
[165] Fia (n 6) 187.
[166] Kerber (n 8) 122.
[167] Purtova (n 64) 18.
[168] ibid.
[169] {Citation}
[170] Commission, 2020a. A European Strategy for Data COM (2020) 66 final.
[171] See DA, Explanatory Memorandum, 13, and recital 5.

Within this framework, one may think that some pieces of legislation may already provide tools for surmounting these barriers. Several laws indeed allow (or impose, in some cases) access to non-personal data held by another entity. However, they have little relevance as a way of boosting data availability in the data economy. This is the other side of the coin – usually neglected by the literature – of the ownership-centric management system of non-personal data[172]

To summarise, the status quo requires adoption of measures aimed at unleashing the economic potential of data. However, the potential granting of property rights in data 'would produce a new scheme of entitlements that is substantively similar to what already exists, thus perpetuating the same frustrations all sides have felt with the existing federal regulations'[173]. Therefore, an open access approach to data has to be implemented. At the same time, both the economic characteristics of data and the risks which such an approach hides should be evaluated. The focus of the legislator should be on the creation of conditions for coopetition in the distribution of data and on the empowerment of the weaker actors in the system – users of IoT products and third parties (SMEs but not only) which need access to data in order to participate effectively on the various markets for goods and services.

Following the findings from this chapter in the next part, the political economy of data under DA is presented, in order to be provided an answer to the question whether the proposed construction of legal relations is capable of remedying the existing shortcomings regarding to data.

[172] Fia (n 6) 187.
[173] Barbara J Evans, 'MUCH ADO ABOUT DATA OWNERSHIP' 25 75.

III. The Impact of the Data Act on the Status Quo

III.1. Context

In retrospect[174], acknowledging the challenges which the European data economy had faced, in 2014 the EC started to consider the adoption of legislative and non-legislative measures to stimulate it by promoting access and reuse of data[175]. Following that, in 2016, the Commission for the first time announced its intention to launch an 'European Free Flow of Data Initiative'[176] and identified obstacles that hampered the development of the nascent European data economy. The Communication on 'Building a European Data Economy'[177] from 2017 is an emblematic next step in the policy development and the first explicit identification of the problem that non-personal data are not reused and shared enough (especially for innovation) as an important policy issue[178]. After the second Communication entitled 'Towards Building a European Data Space'[179] the first pieces of legislation on access and reuse of public[180] and private[181] sector data have been adopted at EU level. When it comes to private sector non-personal data it is to be noted that the approach of the legislator was initially based on the understanding that B2B relations regarding the access to data could be regulated by self-regulatory codes of conduct, that is legally non-binding instruments[182], which has been proven to be insufficient[183].

[174] For more detailed analysis of the evolution of the policies and legislative initiatives for the data economy in the EU see Schmidt-Kessen (n 84) 61–68.
[175] EC (2012). Communication Towards a thriving data-driven economy, COM(2014) 442 final.
[176] Inception Impact Assessment by DG CNECT from November 2016.
[177] EC (2017). Communication Building a European Data Economy COM (2017) 9 final.
[178] Wolfgang Kerber, 'Governance of IoT Data: Why the EU Data Act Will Not Fulfill Its Objectives (Second Version)' (18 July 2022) 121 <https://papers.ssrn.com/abstract=4080436> accessed 19 April 2023.
[179] EC (2018). Communication Towards a Common European Data Space, COM (2018) 232 final.
[180] Directive 2003/98 on the re-use of public sector information, now replaced by Directive (EU) 2019/1024 on open data and the re-use of public sector information.
[181] Regulation 2018/1807 on the free flow of non-personal data in the EU.
[182] Regulation 2018/1807, Art. 6
[183] Zufall and Zingg (n 37) 227.

III.2. Main Objectives, Principles and Instruments of the Data Act

Recognizing that the previous steps have not resulted in fair distribution of data resource and unleash of its potential for generating wealth, in 2019 the Commission took over the agenda on the European data economy and in 2020 it announced its outlook in the Communication 'A European Strategy for Data'[184]. This Strategy restated that 'data should be available to all' because 'this will help society to get the most out of innovation and competition and ensure that everyone benefits from a digital dividend'. The Commission declared also that 'digital Europe should reflect the best of Europe - open, fair, diverse, democratic, and confident'.

One of the central pillars of the strategy is the enactment of a general piece of legislation in the form of a DA[185] to incentivize horizontal, cross-sector data sharing.

The Commission has acknowledged that 'B2B data sharing 'has not taken off at sufficient scale' and has identified that the main causes for this result are 'lack of economic incentives (including the fear of losing a competitive edge), lack of trust between economic operators […], imbalances in negotiating power, fear of misappropriation of the data by third parties, and a lack of legal clarity on who can do what with the data'[186]. As a result, the manufacturers of IoT devices prefer to establish, most often through technical measure exclusive, de facto control over data generated by the IoT devices of the users. As a consequence, neither the users of these IoT devices nor other firms can get enough access to this data[187].

[184] EU Commission (2020). A European Strategy for Data COM(2020) 66 final.
[185] Other core elements of the strategy are the Data Governance Act (DGA), the Public Sector Information (PSI)/Open Data Directive (ODD), and the Regulation on the Free Flow of Non-Personal Data (FFNPDR). Additional initiatives designed to regulate digital services (the Digital Services Act or DSA), digital markets (Digital Markets Act or DMA), artificial intelligence (AI Act), the extraction of informational value from protected works (CSDM) and the processing of personal data (GDPR) could also be perceived as building blocks of the EU Data Governance framework.
[186] Communication 'A European strategy for data'.
[187] DA, Explanatory Memorandum, 13, and recital 5.

Against this backdrop, through the DA the Commission aims to achieve 'fairness in the allocation of value from data among actors in the data economy and to foster access to and use of data'[188]. Whereas, although at first sight this objective appears to be promising, it does not assist the understanding of the political economy of data underlying the DA, because the proposal does not discuss or explain the meaning of this goal at all. This approach is problematic not only because it could make the application of teleological method of interpretation to the act difficult, but also because it indicates that the legislator does not have though understanding of what the fair distribution of data resources across society should be.

Furthermore, it is crucial that to fulfil its objective the proposal does not attempt to question the strategy of the manufacturers to capture the data in an exclusive way through the technical design of their IoT devices[189]. Instead, it tries to limit the ensuing negative effects through two instruments: (1) Introduction of non-waivable user rights (Art. 4 and 5), and (2) Introduction of contractual agreements between data holders and users about whether and how the data holders can use the non-personal IoT data (Art. 4(6) s.1). It is important that the novel user rights include rights of the users to share their IoT data with a third party (TP), which imposes conclusion of a negotiated agreement between the data holder and the TP about the conditions under which the TP can use this IoT data and this contract can be interpreted as a 'licensing agreement'[190].

The mechanism adopted by the DA is heavily criticized as weak and largely ineffective[191]. Furthermore, the proposal has come under criticism from some Member States[192] which

[188] DA, Explanatory Memorandum, 2.
[189] Eckardt and Kerber (n 47) 9.
[190] Kerber (n 5) 54.
[191] Eckardt and Kerber (n 47) 15–18; Kerber (n 8); Inge Graef and Martin Husovec, 'Seven Things to Improve in the Data Act' (7 March 2022) <https://papers.ssrn.com/abstract=4051793> accessed 16 June 2023.
[192] The Netherlands (Ministerie van Buitenlandse Zaken, 'Non-Paper on the Data Act - Publication - The Netherlands at International Organisations' (1 October 2021) <https://www.permanentrepresentations.nl/documents/publications/2021/10/1/non-paper-on-the-data-act> accessed 29 May 2023.) but also Denmark.

support the development of market-based mechanisms, including the development of common, private-led data sharing standards, instead of binding data transfers obligations for businesses[193]. Even though I recognize that the way in which the specific rights and obligations are formulated and legislatively guaranteed could be in practice as important as the granting of the rights itself, I focus in my analysis more on the concepts implemented by the legislator in the act, than on the wording, the structure and the operational deficits of the proposal. This is because, on the one hand, several publications have already reacted to the formulation of the rights and obligations under the DA[194], and on the other hand, since the proposal seems to be more or less ambivalent regarding its understanding for fair allocation of data value and different ideas have been proposed in the literature, there remains room for discussions regarding the way in which political economy of non-personal data should be designed.

III.3. Scope Ratione Materiae of the Data Act

The problem for the scope rationae materiae of the DA, namely which data are governed by the act, is particularly relevant for two reasons – first, it should be noted that all non-personal data with no legal status which are not covered by the regulation at stake will continue to be governed by the non-legal rules of the de facto ownership regime; second, too narrow scope of the relevant data set to which the act provides access would make the proposal's contribution

[193] Clément Perarnaud and Rosanna Fanni, 'Towards a New European Data Revolution?' 2.
[194] Lord John Thomas and others, 'Response of the European Law Institute to the Public Consultation on a Data Act' (2021) <https://papers.ssrn.com/abstract=4096872> accessed 19 April 2023; Markus Lampinen and Paulius Jurcys, 'Prifina Comments on the Regulation of the European Parliament and of the Council on Harmonised Rules on Fair Access to and Use of Data (Data Act)' (13 May 2022) <https://papers.ssrn.com/abstract=4110462> accessed 16 June 2023; Moritz Hennemann and Gregor Lienemann, 'The Data Act - Article-by-Article Synopsis of the Commission Proposal' (1 March 2022) <https://papers.ssrn.com/abstract=4079615> accessed 16 June 2023; Beatriz Botero Arcila and Teodora Groza, 'Comments to the Data Act from the Law and Technology Group of Sciences Po Law School' (13 June 2022) <https://papers.ssrn.com/abstract=4135212> accessed 16 June 2023; Ducuing and others (n 23); Drexl and others (n 48); Can Atik, 'Data Act: Legal Implications for the Digital Agriculture Sector' (23 June 2022) <https://papers.ssrn.com/abstract=4144737> accessed 16 June 2023.

to its ambitious goals of unlocking the value of data in Europe and enhancing opportunities of innovation[195] rather nebulous.

It must be clear that if all limitations laid down in the proposal are maintained, the act will not be capable of impacting the currently existing political economy of data. Furthermore, the justifications of some restrictions are dubious.

First, the Act only applies to IoT data and it does not cover online service-related usage data, which undoubtedly hide great latent potential. Scholars identify three main arguments in favour of this restriction - first, these data will often qualify as 'personal data'; second, for many types of behavioural service-related data there will be substitutes and therefore, sometimes, access to the individual-level data controlled by one specific data holder may be less important compared to access to product usage data; third, online service providers frequently chose business models that rely on the monetization of data[196]. However, none of the arguments put forward is uncontested – first, the data act aims to regulate the access to both personal and non-personal data (some of the IoT data will be also personal); second, in fact, the value of individual-level data is limited but this is valid for the IoT data too; third, legitimate interests of service providers must be taken in consideration, however, it is feasible the balance between the interests of the different parties to be found, e.g., third parties might be obliged to pay for the data access and non-compete clauses similar to those set out in Arts. 4(4) and 6(2)(e) DA could be adopted.

Second, the proposal requires a making available of only 'data generated by the use of a product or related service'[197]. However, it does not define what data can be considered as 'generated'

[195] See Explanatory Memorandum, DA, 1.
[196] Heike Schweitzer and Axel Metzger, 'Data Access Under the Draft Data Act, Competition Law and the DMA: Opening the Data Treasures for Competition and Innovation?' (22 December 2022) 348 <https://papers.ssrn.com/abstract=4309694> accessed 19 April 2023.
[197] Art. 3(1), 4(1) and 5(1)

– from Recital 31 it is known that the act applies to 'actively provided' and 'passively observed' data, i.e., only raw data. Still, in order for third parties to provide additional services to users, such as repair or predictive maintenance services in downstream or adjacent markets of other new goods or IoT-related services, it may be necessary they to be granted with access to inferred and derived data[198]. Since, for many of these services, it is not sufficient to have only access to raw data, the full exclusion of these types of data can lead to a data set that might be much too narrow to enable third parties actively participate on the market[199]. Furthermore, the data sets are often mixed and this limitation could be used as an excuse for rejected access to such sets.

The question for raw and generated data rises debates[200]. Some authors advocate for the replacement of the conduct-based approach with a purposed-based approach, i.e., data access and use rights to depend on the existence of legitimate interest of a user, even if the user has not contributed so much to data generation, and in a case of a minor interest and lower level of dependence, the contribution to the generation of the data has to be larger[201]. Defenders of this suggestion clarify that the suggested approach presuppose delineating the data more narrowly, namely, only with regard to the data based on the first encoding that follows the use of the product or the related service[202].

This suggestion is worthy of serious discussion. From the analysis of data as a resource it is apparent that raw data often have limited economic importance especially when it is further limited to individual-level data. Therefore, the total exclusion of the inferred and derived data from regulation would mean that the allocation of wealth of these types of data would be subject

[198] Kerber (n 8) 127.
[199] ibid.
[200] Drexl and others (n 48) 10–15.
[201] ibid 13.
[202] ibid 14.

to non-legal rules which could retain the existing legal uncertainty. Here, the legislator should find the balance between the interests of the IoT users who undoubtedly contribute for the production of these data (even if it is to limited extent) and paid price for the embedded in the IoT products software which often processes data on the first level, the interests of the manufactures of IoT products and investors in analytical software whose investments must be guaranteed, the interests of the third parties who can use these data for developing goods and services, and the general interest of the society in effective competition and vigorous innovation process.

In light of the above I find the proposal the data based on the first encoding that follows the use of the product or the related service to be covered by the rights granted by the DA reasonable. At the same time, an introduction of an obligation for payment of a fee, when the data at stake have been processed by a software for which the user has not paid, seems justified.

The third limitation which is discussed in more details in the following sub-chapters stems from the fact that the proposal regulate only access to individual-level data, which cap the practical relevance of the DA.

The above analysis could not give an overall answer to the very big question how should the access to the various types of data be regulated. The mentioned problems require in-depth examination from many perspectives (legal, economic, data science, etc.). However, the chapter aims to demonstrate that taken together the limitations would exclude a significant amount of data sets and connected solutions from the scope of the Regulation[203] which could prejudice the achievement of the sensible political economy objectives of the legislator. Simultaneously, the paper aims to demonstrate that various legislative solutions could be discussed in order the scope ratione materiae of the act to be expanded. Even though the wide

[203] Atik (n 194) 12.

variety of data types could humper the regulation of the data governance, it cannot justify exclusion of valuable resources from the rules which allocate wealth among the society, because the absence of regulatory intervention when it comes to data means data enclosure and de facto exclusive economic rights.

It should be recalled here the case with personal data and conclusion which Nadezhda Purtova has reached that 'the core question of data protection should be not if there should be property rights in personal data, but how to structure those property rights in a way that is both economically sound and respectful of the individual's rights'[204]. Similarly, the question of non-personal data should be how to structure the rights in a way that is both economically sound and respectful of the legitimate interests of the actors within the data economy.

III.4. Status of the IoT Users, Data Holders and Third Parties

In the light of the above discussions this chapter analyses how the wealth of data is distributed among the actors within the data economy. It must be noted first of all that, potentially weak position of the IoT users and the third parties would amount to maintenance of the status quo, i.e., enclosure and prioritization of non-personal data by the data holders. However, risk of enclosures can emerge from other hypotheses too.

III.4.1 *Status of the IoT Users*

Interestingly enough, the DA puts in the centre of the data economy the user of IoT. To achieve its objectives the legislator relies on mechanism based on the introduction of new non-waivable rights of the users to access and share the data they have generated through their IoT devices

[204] Purtova (n 64) 29.

(Arts. 4-5). In addition, Article 4(6) stipulates that the 'data holder shall only use any non-personal data generated by the use of a product or related service on the basis of a contractual agreement with the user'.

This apparatus deserve critique because it runs the risk of perpetuating the data enclosure which currently exists – if the users rights proven weak, they will not be able to overcome the process of propertization of non-personal data in favour of the data holders. If, however, these rights happen to be strong they could lead to the situation known as tragedy of the data commons. In addition, this mechanism does not reflect the established understanding, recognized also in the ALI-ELI Principles[205], that the co-generators should have rights in data. But what is the most important, in any case this model cannot open significantly the access to data and unleash their economic potential.

Scholars[206] classify the proposed access right and data portability right as statute-based rights in personam (not in rem); they both are only effective against data holders and against providers of data processing services, respectively, i.e., they are not rights erga omnes. At the same time, it should be noted that the rights resemble IP-akin type[207] of a data producer's right, because they may be interpreted as containing a component of exclusivity – Art. 4 (6)[208].

The contract between the user and data holder is seen as 'the elephant in the room' of the DA, because, on the one hand, the provision of Art. 4 (6) is theoretically a big step for the empowerment of consumers with respect to their IoT data, because without their consent the data holders cannot use them, but on the other hand, the act does nearly nothing to support the

[205] Principles 18-19.
[206] Charlotte Ducuing, 'An Analysis of IoT Data Regulation under the Data Act Proposal through Property Law Lenses' (20 September 2022) 6 <https://papers.ssrn.com/abstract=4225027> accessed 19 April 2023; Geiregat (n 34) 43.
[207] For IP-akin rights see Schmidt-Kessen (n 84) 81.
[208] From Art 35 some authors extract arguments against the understanding that Art. 4(6) regulates an exclusive right, but they recognize that the provision could be misunderstood in practice (Drexl and others (n 48) 19.), therefore the hypothesis should be discussed.

consumers use this theoretically strong position for exercising more control over their IoT data[209]. If this is the case if follows that the legislator relies on a 'paper tiger' to guarantee the fair allocation of data.

If, on the contrary, users tend to take advantage of their empowerment they can create a new barrier before the free flow of non-personal data. In case in which the users prefer to reject access to their data this will result in tragedy of data commons.

A hypothesis of tragedy of data anti-commons is also possible – the act provides that the user is the central right-holder. However, if she is neglecting to employ the data, this can result in the resource not being used at all.

Both hypotheses which are equally jeopardizing the open access to data. They are both a possible result of the decision to 'put the users in the driver's seat'[210], which does not take into account that 'the context is too complex, information to make informed choices scarce, the incentives unclear and individuals can be easily nudged in different directions and/or are too busy managing their own lives'[211].

The greatest weakness of this model, however, is that it does not provide a legal remedy for access to aggregated datasets, because a user can grant third party only access to her own data, which does not allow open access to big volumes of data, which are much more valuable.

III.*4.2. Status of the Data Holders*

When it comes to the data holders it should be noted that the Commission has recognized that the manufacturer of IoT devices typically has exclusive control over the use of data generated

[209] Kerber (n 8) 132.
[210] Geiregat (n 34) 23.
[211] Botero Arcila and Groza (n 194) 3.

by the use of a product or related service, which contributes to lock-in effects and hinders market entry for players offering aftermarket services[212]. Further, the Commission has declared preferences for non-exclusive rights[213].

However, some scholars[214] argue that the DA will actually strengthen the position of data holders and claim that the legislator has failed to find the proper balance between the need to ensure the incentives for generating and collecting IoT data and extracting value from them and the declared objectives of the regulation. Others[215] identify various imperfections of the way in which the rights and obligations of the data holders are formulated and ensured. These weaknesses should be considered by the legislator because they can potentially compromise the access to data in practice. On conceptual level, however, I find worrying that the material scope of the Regulation is too narrow to challenge the position of exclusive control which the data holders de facto have.

III.4.*3. Status of the Third Parties*

Under the proposal third parties have a 'derived' right of access only: they may process the data 'only for the purposes and under the conditions agreed with the user' (Art. 6(1)). As I have already mentioned, in my opinion, the greatest weakness of the proposal on the conceptual level is the omission to grant a direct right to access to aggregated non-personal data sets to third parties who have had no part in the generation of these data.

Three main reasons require this legislative decision to be seriously reviewed and reconsidered – first, the objectives which the EC has outlined presuppose wider access to non-personal data,

[212] Explanatory Memorandum, DA.
[213] Recital 6
[214] Kerber (n 5).
[215] Drexl and others (n 48) 32–44.

second, the needs of the market, especially regarding the development of the AI technologies, call for such a step, third, the economic characteristics of data as a resource[216].

While the Communication 'Building a European Data Economy' clearly states as key objective facilitation and promotion of access for more players to larger and more diverse data sets, which would allow for a higher rate of innovation and the development of data-intensive technologies, it can be no doubt that the derived right to access to individual level raw data cannot be an adequate solution. Furthermore, an ever-growing demand for access to big datasets[217] lives side by side with the continuing reluctance of the data holders to share voluntarily[218], while the competition law offers limited (if any) remedies for the access seekers[219].

In light of the above, it seems that the DA cannot effectively challenge the enclosure of non-personal data and the consecutive under-use of these resources without empowering the third parties to directly access big sets of non-personal data. Of course, such a right should be balanced with the interests of the data holders, which should be incentivised to collect, store and transfer in interoperable format the generated data.

[216] Infrastructural, non-rivalrous, non-depletable resource, which acquisition and usage are lossless; characterized by dynamic economies of scale and economies of scope, which create data-driven indirect network effects; cheap to be reproduced and easy to be shared;
[217] 'The Right to Process Data for Machine Learning Purposes in the EU by Mauritz Kop :: SSRN' (n 104).
[218] Which is obvious from the failure ('limited efficacy') of the soft law approach under the Free-Flow of Non-Personal Data Regulation - Ducuing and others (n 23) 59.
[219] Schweitzer and Metzger (n 196) 348; Schmidt-Kessen (n 84).

CONCLUSION

It is clear that within the data economy data constitute a resource. Therefore, regulating the right to access to data (right which in the context of this specific resource replaces the classic usus-fructus right) the DA represents an instrument for allocation of wealth. The proposal declares that its objective is to provide 'fairness in the allocation of value from data among actors in the data economy'. It should be clear, however, that the mere reference to fairness should not be used to justify regulations that pursue distributive purposes without the ability to rely on explicit extra value judgments and aims in the law. Therefore, when building such a new framework, the legislator, as a first step, must provide a comprehensive analysis of the object of regulation and the existing de facto relations regarding this object. In the case at stake the Commission has on several occasions correctly identified the enclosure of non-personal data by the data holders as the central problem within the data economy. However, the combination of strong objectives and very limited material scope makes the legislative proposal nebulous. Some of the shortcomings are probably due to misunderstanding of the characteristics of the regulated object (e.g., the fact that some types of valuable data are left without legal status), incorrect assessment of the roles of the market players (e.g., entrust of the users to serve as 'traffic cops' for the allocation of access to data), misappraisal of the needs on the market (e.g., the need for access to Big Data for the AI industry). In light of the above, it is obvious that a data economy is always a complex distributive game in which a balance between numerous (legitimate) interests should be found. It must be clear, there is no one fit all solution for all types of non-personal data and for all types of relations (B2B, B2C, B2G), but the complexity of the subject matter does not justify deny of regulation. When limits the scope of the proposal, the Commission jeopardizes the objectives of its Strategy for Data and risks to maintain the status qua without a substantive change.

BIBLIOGRAPHY

Atik C, 'Data Act: Legal Implications for the Digital Agriculture Sector' (23 June 2022) <https://papers.ssrn.com/abstract=4144737> accessed 16 June 2023

Botta M 'Shall We Share? The Principle of FRAND in B2B Data Sharing:: SSRN' <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4423805> accessed 21 May 2023

Botero Arcila B and Groza T, 'Comments to the Data Act from the Law and Technology Group of Sciences Po Law School' (13 June 2022) <https://papers.ssrn.com/abstract=4135212> accessed 16 June 2023

Drexl J and others, 'Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act)' (25 May 2022) <https://papers.ssrn.com/abstract=4136484> accessed 11 June 2023

Ducuing C, 'An Analysis of IoT Data Regulation under the Data Act Proposal through Property Law Lenses' (20 September 2022) <https://papers.ssrn.com/abstract=4225027> accessed 19 April 2023

Ducuing C and others,'White Paper on the Data Act Proposal' (26 October 2022) <https://papers.ssrn.com/abstract=4259428> accessed 16 June 2023

Eckardt M and Kerber W, 'Property Rights Theory, Bundles of Rights on IoT Data, and the Data Act' (26 February 2023) <https://papers.ssrn.com/abstract=4376833> accessed 21 May 2023

Erp S van and Swinnen K, 'The Legal Status of Co-Generated Data: With Particular Focus on the ALI-ELI Principles for a Data Economy and the Rules on Accession, Commingling and Specification' (2022) 2022 Technology and Regulation 61

Evans BJ, 'MUCH ADO ABOUT DATA OWNERSHIP' 25

Fia T, 'An Alternative to Data Ownership: Managing Access to Non-Personal Data through the Commons' (2021) 21 Global Jurist 181

Geiregat S, 'The Data Act: Start of a New Era for Data Ownership?' (8 September 2022) <https://papers.ssrn.com/abstract=4214704> accessed 13 June 2023

Graef I, Gellert R and Husovec M, 'Towards a Holistic Regulatory Approach for the European Data Economy: Why the Illusive Notion of Non-Personal Data Is Counterproductive to Data Innovation' (27 September 2018) <https://papers.ssrn.com/abstract=3256189> accessed 8 June 2023

Graef I and Husovec M, 'Seven Things to Improve in the Data Act' (7 March 2022) <https://papers.ssrn.com/abstract=4051793> accessed 16 June 2023

Hennemann M and Lienemann G, 'The Data Act - Article-by-Article Synopsis of the Commission Proposal' (1 March 2022) <https://papers.ssrn.com/abstract=4079615> accessed 16 June 2023

Hummel P, Braun M and Dabrock P, 'Own Data? Ethical Reflections on Data Ownership' (2021) 34 Philosophy & Technology 545

Kaplow L and Shavell S, 'Economic Analysis of Law' (1 February 1999) <https://papers.ssrn.com/abstract=150860> accessed 6 June 2023

Kerber W, 'Governance of IoT Data: Why the EU Data Act Will Not Fulfill Its Objectives (Second Version)' (18 July 2022) <https://papers.ssrn.com/abstract=4080436> accessed 19 April 2023

Lametti D, 'The Concept of the Anticommons: Useful, or Ubiquitous and Unnecessary?' (2 January 2013) <https://papers.ssrn.com/abstract=2195549> accessed 12 June 2023

Lampinen M and Jurcys P, 'Prifina Comments on the Regulation of the European Parliament and of the Council on Harmonised Rules on Fair Access to and Use of Data (Data Act)' (13 May 2022) <https://papers.ssrn.com/abstract=4110462> accessed 16 June 2023

Lee S-Y, Jung S and Hwang WJ, 'ALI-ELI PRINCIPLES FOR A DATA ECONOMY - DATA TRANSACTIONS AND DATA RIGHTS -' (2022) 107 Korea Law Review 493

Martens B and others, *Business-to-Business Data Sharing: An Economic and Legal Analysis* (European Commission 2020)

Mezzanotte F, 'Access to Data: The Role of Consent and the Licensing Scheme' [2017] S. Lohsse, R. Schulze, D. Staudenmayer (eds.), Trading Data in the Digital Economy: Legal Concepts and Tools, Hart/Nomos, Oxford/Baden-Baden <https://www.academia.edu/37037356/Access_to_Data_The_Role_of_Consent_and_the_Licensing_Scheme> accessed 28 May 2023

Opderbeck DW, 'Socially Rivalrous Information: Of Candles, Code, and Virtue' (21 August 2007) <https://papers.ssrn.com/abstract=1008500> accessed 27 May 2023

Park K, 'Data as Public Goods or Private Properties?: A Way Out of Conflict Between Data Protection and Free Speech' (2021) 6 UC Irvine Journal of International, Transnational, and Comparative Law 77

Perarnaud C and Fanni R, 'Towards a New European Data Revolution?'

Picht P G 'Caught in the Acts: Framing Mandatory Data Access Transactions Under the Data Act, Further EU Digital Regulation Acts, and Competition Law by :: SSRN' <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4076842> accessed 19 April 2023

Prüfer J and Schottmüller C 'Competing with Big Data* 2021 - The Journal of Industrial Economics - Wiley Online Library' <https://onlinelibrary.wiley.com/doi/full/10.1111/joie.12259> accessed 27 May 2023

Purtova N, ‘Illusion of Personal Data as No One’s Property’ (29 October 2013) <https://papers.ssrn.com/abstract=2346693> accessed 24 May 2023

Reimsbach-Kounatze C, ‘Enhancing Access to and Sharing of Data: Striking the Balance between Openness and Control over Data’ (2021)

Schmidt-Kessen MJ, ‘The Impact of Data Ownership Rights on Competition in Big Data Markets: Reflections in the Context of the EU and Global Data Economy’ (2020) 2 VIT Law Review 56

Schweitzer H and Metzger A, ‘Data Access Under the Draft Data Act, Competition Law and the DMA: Opening the Data Treasures for Competition and Innovation?’ (22 December 2022) <https://papers.ssrn.com/abstract=4309694> accessed 19 April 2023

Singh PJ and Vipra J, ‘Economic Rights Over Data: A Framework for Community Data Ownership’ (2019) 62 Development 53

Soave T, ‘Digital Humanitarians and International Lawyers: Worlds Apart or Two Sides of the Same Coin?’ (2022) 25 Max Planck Yearbook of United Nations Law Online 718

Somaini L, ‘Regulating the Dynamic Concept of Non-Personal Data in the EU: From Ownership to Portability’ (2020) 6 European Data Protection Law Review (EDPL) 84

Streel A, Krämer J and Senellart P, ‘Making Data Portability More Effective for the Digital Economy’ [2021] SSRN Electronic Journal

Kop M, ‘The Right to Process Data for Machine Learning Purposes in the EU by :: SSRN’ <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3653537> accessed 30 May 2023

Krämer J and others ‘The Role of Data for Digital Markets Contestability: Case Studies and Data Access Remedies’ (*CERRE*) <https://cerre.eu/publications/data-digital-markets-contestability-case-studies-and-data-access-remedies/> accessed 27 May 2023

Thomas LJ and others, ‘Response of the European Law Institute to the Public Consultation on a Data Act’ (2021) <https://papers.ssrn.com/abstract=4096872> accessed 19 April 2023

Thouvenin F, ‘Wem gehören meine Daten? Zu Sinn und Nutzen einer Erweiterung des Eigentumsbegriffs’ (2017) 113 Schweizerische Juristen-Zeitung 21

Thouvenin F and Tamò-Larrieux A, ‘Data Ownership and Data Access Rights: Meaningful Tools for Promoting the European Digital Single Market?’ in Mira Burri (ed), *Big Data and Global Trade Law* (Cambridge University Press 2021) <https://www.cambridge.org/core/books/big-data-and-global-trade-law/data-ownership-and-data-access-rights/BC314C63C58A09C4B9C5D55894FE68C6> accessed 11 June 2023

Tombal T, ‘The Rationale for Compulsory B2B Data Sharing and Its Underlying Balancing Exercises’ (1 December 2021) <https://papers.ssrn.com/abstract=3984873> accessed 21 May 2023

Tombal T and Graef I, 'The Regulation of Access to Personal and Non-Personal Data in the EU: From Bits and Pieces to a System?' (15 December 2022) <https://papers.ssrn.com/abstract=4304148> accessed 21 May 2023

Wilkinson MA and Lokdam H, 'Law and Political Economy' (20 March 2018) <https://papers.ssrn.com/abstract=3144723> accessed 12 June 2023

Yakowitz J, 'Tragedy of the Data Commons' Harvard Journal of Law & Technology Volume 25, Number 1 Fall 201125

Zaken M van B, 'Non-Paper on the Data Act - Publication - The Netherlands at International Organisations' (1 October 2021) <https://www.permanentrepresentations.nl/documents/publications/2021/10/1/non-paper-on-the-data-act> accessed 29 May 2023

Zech H, 'Data as a Tradeable Commodity' in Alberto De Franceschi (ed), *European Contract Law and the Digital Single Market: The Implications of the Digital Revolution* (Intersentia 2016) <https://www.cambridge.org/core/books/european-contract-law-and-the-digital-single-market/data-as-a-tradeable-commodity/9D9E07D5B5E3C86C1B2E04BF8C01D15B> accessed 20 May 2023

Zimmer D, 'Property Rights Regarding Data?', *Trading Data in the Digital Economy: Legal Concepts and Tools, Münster Colloquia on EU Law and the Digital Economy III* (2017)

Zufall F and Zingg R, 'Data Portability in a Data-Driven World' in Ching-Fu Lin, Shin-yi Peng and Thomas Streinz (eds), *Artificial Intelligence and International Economic Law: Disruption, Regulation, and Reconfiguration* (Cambridge University Press 2021) <https://www.cambridge.org/core/books/artificial-intelligence-and-international-economic-law/data-portability-in-a-datadriven-world/F445EC4A9E9665A05E773A88E8840027> accessed 20 May 2023