MOBIDATALAB

Labs for prototyping future mobility data sharing solutions in the cloud

D2.7 Data Governance Assesment

23/01/2023

Author(s): Emre BAYAMLIOĞLU (KUL), Aliki BENMAYOR (KUL)

MobiDataLab is funded by the EU under the H2020 Research and Innovation Programme (grant agreement No 101006879).

Summary sheet

Deliverable Number	D2.7
Deliverable Name	Data Governance Assessment
Full Project Title	MobiDataLab, Labs for prototyping future Mobility Data sharing cloud solutions
Responsible Author(s)	Emre BAYAMLIOĞLU (KUL), Aliki BENMAYOR (KUL)
Contributing Partner(s)	-
Peer Review	CNR, AKKA
Contractual Delivery Date	30-04-2022
Actual Delivery Date	29-04-2022
Status	Final
Dissemination level	Public
Version	V1.0
No. of Pages	63
WP/Task related to the deliverable	WP2 / T2.5
WP/Task responsible	AKKA / KUL
Document ID	MobiDataLab-D2.7-DataGovernanceAssessment -v1.0
Abstract	This deliverable seeks to explain the constituent elements of data governance and sets out the different data governance models. Based on the use cases provided in D2.9 of the project, it explores how different types of data governance models and mechanisms (as collaborative data empowerment settings) could accommodate and support data access and sharing in the Transport Cloud. It also covers the legal entitlements on data that affect data transactions.

Legal Disclaimer

MOBIDATALAB (Grant Agreement No 101006879) is a Research and Innovation Actions project funded by the EU Framework Programme for Research and Innovation Horizon 2020. This document contains information on MOBIDATALAB core activities, findings, and outcomes. The content of this publication is the sole responsibility of the MOBIDATALAB consortium and cannot be considered to reflect the views of the European Commission.

Project partners

Organisation	Country	Abbreviation
AKKA I&S	France	AKKA
CONSORZIO INTERUNIVERSITARIO PER L'OTTIMIZZAZIONE E LA RICERCA OPERATIVA	Italy	ICOOR
AETHON SYMVOULI MICHANIKI MONOPROSOPI IKE	Greece	AETHON
CONSIGLIO NAZIONALE DELLE RICERCHE	Italy	CNR
HOVE	France	HOVE
HERE GLOBAL B.V.	Netherlands	HERE
KATHOLIEKE UNIVERSITEIT LEUVEN	Belgium	KUL
UNIVERSITAT ROVIRA I VIRGILI	Spain	URV
POLIS - PROMOTION OF OPERATIONAL LINKS WITH INTEGRATED SERVICES	Belgium	POLIS
F6S NETWORK IRELAND LIMITED	Ireland	F6S

Document history

Version	Date	Organisation	Main area of changes	Comments
0.1	17/03/2022	KUL	All	Draft for review
0.2	1/04/2022	AKKA, CNR	All	Peer review
0.3	6/04/2022	Stephane Dreher (OXSYS)	All	Advisory Board Review
0.4	21/4/2022	KUL	All	End of rework phase
0.5	29/04/2022	AKKA	All	End of Quality check
1.0	29/04/2022	KUL - AKKA	All	Final version

Funded by the European Union

Executive Summary

Data-driven technologies are considered the key driver of the current and future commercial and other organisational initiatives and activities. With the exponential increase in the volume and diversity of data, entities need more advanced and innovative approaches to combine, manipulate, store, and extract value from data. In this respect, stakeholders of the data economy are increasingly confronted with several strategic requirements such as legal compliance and maximizing value creation. This brings to fore *data governance* as an emerging concept denoting a system of *decision rights, roles* and *accountabilities* for organisational data use and data-driven processes.

In line with this, the project MobiDataLab aims to propose to the mobility stakeholders (transport organising authorities, operators, industry, innovators) a methodology and tools to foster the development of fair data access and sharing regime**.** The Project leverages the legal, technological and economic opportunities to guide on how to improve the quality, accessibility and usability of the mobility and transport data. This document, Deliverable 2.7 (D2.7) under WP2 of the Project, provides an assessment of data governance models/mechanisms and analyses legal frameworks that affect data transactions/contracts as an essential pillar of data governance.

Chapter II primarily explains the concept of data governance from an interdisciplinary perspective and lays out a conceptual framework comprising: i) *organisational*, ii) *legal and regulatory,* and iii) *technological* dimensions. The framework enables a more systematic approach to the organisational and transactional necessities of data governance within a broader picture encompassing the whole regulatory landscape on personal and nonpersonal data. The Chapter further provides an account of suggested forms of data governance models and mechanisms as referred to in various sectoral initiatives, legislative proposals and research projects together. The aim is to identify the variables that have an impact on the data governance mechanisms together with the actions that are required to allow such data governance mechanisms. The Chapter also includes an analysis of the essential features of these models/mechanisms and explores their suitability for the Transport Cloud *use cases* as real-life data governance scenarios. The findings of the Chapter reveal the lack of consensus as to the exact nature, characteristics and legal status of these organisational structures. It becomes clear in real-life scenarios illustrated by the use cases that there exists no mature enough market structures or business practices in the mobility or transport sector which could be easily associated with a model or mechanism with precise features and uniform application. All models and mechanisms present similar legal challenges that can partly be linked to uncertainties about privacy, permissible types of data use and technical implementation which act as a deterrent for researchers, investors and initiators. Considering the complexity and the dynamism of the European legal landscape relating to data governance, it could be concluded that there is a need for further refinement and theorising regarding these models and mechanisms.

Despite the lack of a default legal status, there exist several laws providing substantive rights which interfere with and affect data transactions/contracts in various aspects. Chapter III, provides an analysis of the relevant legal regimes (i.e., personal data protection, copyright, *sui generis* database right and trade secret protection) that contain the substantive rights applicable to contracts aiming to make available both personal and non-personal data. First, where data can be linked to an identified or identifiable natural person, the rules on personal data protection (GDPR) limit the circulation and

accessibility of data. This being said, the GDPR also provide initiatives through individual rights such as the right to data portability, the right to erasure and the right to revocation of consent data subjects' rights contribute to a trusted environment for the data subjects. Considering that data transactions involving personal data heavily rely on consent, the analysis reveals that consent makes up the weak spot of data transactions simply for the reason that under the GDPR, withdrawal of consent is an inalienable and irrevocable fundamental right. In sum, there is a need for a restrictive legal interpretation of the right to withdraw consent for a more stable and foreseeable legal environment for data transactions involving personal data. *Second*, IP rights lay out exclusive rights on certain forms of intangible elements (including data), and thus provide the legal basis for a variety of data transactions. Individual data items (datum) are subject to copyright protection as they amount to original creations such as user-generated text, video or images. More importantly, databases are protected by the *sui generis right* of the Database Directive. The relevant parts provide a comprehensive analysis of the applicability of *sui generis* protection to databases and highlight the relevant discussions, including the case law of the Court of Justice of the European Union (CJEU) on certain unresolved questions. *Third*, trade secret (TS) protection is another legal regime that confers control over data and thus could be of relevance to data transactions/contracts. While the legal protection of sensitive business information could foster data transactions by unlocking data that would otherwise remain undisclosed, it is also possible that overreliance on TS protection could severely impede the efficient operation of data markets. In sum, these legal regimes (i.e., personal data protection, IP rights and TS protection) by no means make up the entire legal landscape that could fully determine the legal limits of data transactions. The analysis rather provides a macro view of the legal frameworks which lay out the substantive rights underlying data transactions.

Chapter IV focuses on the Data Act proposal as an essential building block of the emerging EU data governance regime as laid out in the 2020 European Data Strategy. The proposal which was released on 23.02.2022 is expected to play a key role in the digital transformation in line with the 2030 digital objectives. Complementing the Data Governance Act, the Data Act provides a data access right for the users of the IoT devices, general rules relating to obligations to make data available, a fairness test applicable to data transactions/contracts and a legal framework for the release of private data to public sector bodies. The Act also contain provisions that simplify switching between the cloud service providers and increase interoperability. Addressing the uncertainties (discussed in Ch. III) about the application of *sui generis right*, the Act provides that the sui generis right does not apply to machine-generated data subject to the data access right. Regarding the impact of the Data Act on the MobiDataLab Project, the analysis reveals that the proposed Act will significantly impact the cloud market and current contractual framework and thus the activities and services contemplated within the project. The data access right as being the most concrete intervention together with the exemption of *sui generis right* will be directly applicable to the mobility devices connected to the internet. The general rules relating to the obligations to make data available will be directly applicable to data accessibility rules which will be introduced by the Delegated Regulations within the framework established by the Intelligent Transport Systems Directive.

Table of contents

Funded by the European Union

Abbreviations and acronyms

Abbreviation	Meaning
AI	Artificial Intelligence
API	Application Programming Interfaces
CJEU	Court of Justice of the EU
DA	Data Act
DGA	Data Governance Act [Proposal]
DMA	Digital Markets Act
DSA	Digital Services Act
DSS	Data Sharing Services
EC	European Commission
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor

Funded by the European Union

EEA	European Economic Area
EU	European Union
GDPR	General Data Protection Regulation
IoT	Internet of Things
IP	Intellectual Property
STS	Science and Technology Studies
TFEU	Treaty on the Functioning of the European Union
TS	Trade Secret
WTO	World Trade Organization

1. Introduction

1.1 Project overview

Digitisation and data-driven services in all modes of transport (passenger and freight) are essential enablers for the transformation to safer, more efficient, accessible and sustainable mobility.[1] In line with this, the MobiDataLab project (the "Project") aims to propose to the mobility stakeholders (transport organising authorities, operators, industry, innovators) a methodology and tools that foster the development of a fair data sharing and data access ecosystem**.** The Project leverages the legal, technological and economic opportunities to guide on how to improve the quality, accessibility and usability of the mobility and transport data and bring together mobility stakeholders to find innovative solutions to concrete challenges in terms of making data available to wider use.

MobiDataLab adopts an approach of continuous co-development of knowledge and technical solutions for making data available involving all data producers and consumers in the transport and mobility landscape. The Open Knowledge Base (WP2) is envisaged as a web-based tool for exploiting a large set of data resources and is intended to serve as a reference for practices and solutions responding to a variety of issues such as *interoperability, the applicable legal frameworks, privacy, licensing, and data governance mechanisms.* On the technological front, the Project develops a cloud-based prototype platform (Transport Cloud) which is envisaged as an agnostic model that could be deployed both in B2B and B2G contexts. The Transport Cloud is intended as a unique point of access to public transport, road and mobility data in selected areas which will be made available through a standardized reference data catalogue. Additional datasets may also be included to enrich transport datasets (e.g., environmental data from sensors, weather, pollution, satellite imagery and other open data).

Based on federated cloud principles tested and proven in EU-funded projects, it integrates access and interaction with internal and external resources in a one-stop interface. The Transport Cloud primarily aims to develop solutions to address the technical limitations identified as barriers to data reuse. Given that the stakeholders of a data eco-system have different motivations and concerns,

[1] European Commission, 'Smart and Sustainable Mobility Strategy', 9 December 2020, COM/2020/789 final.

the respective roles and powers must be mapped to provide input for planning and designing the MobiDataLab cloud prototype that meets the expectations.

MobiDataLab deploys several use cases involving different stakeholders of the transport chain and illustrating existing initiatives, products and services in the data sharing and reuse process. This is expected to provide solutions in the form of digitalised services and data sharing and access initiatives with environmental benefits and economic gains. Through use cases representing different aspects of real-life transport situations, MobiDataLab will clarify which types of products and services would better serve the ends to overcome inefficiencies in transport systems by assessing the project use cases.

Taking stock of the above, certain issues regarding the legal implications of the Project and the Transport Cloud come to the fore. The first and most significant challenge is the patchwork of laws that give partial entitlements to data. In this fragmented regulatory landscape, there exist no general "rights" or "entitlements" on data that could be transferred to permit the exclusive use of data by others. Second, the upcoming legislative agenda laid out by the 2020 European Data Strategy regarding data access and sharing is a continuously evolving domain and present a complex web of rules. The two legislative proposals on a Data Governance Act ("DGA") and Data Act ("DA") constitute the main pillars of the emerging EU data governance regime dealing with several aspects of data governance. Third, the contractual matters regarding data access and sharing present a complex nature due to the diversity of the types and data uses, technical tools and the legal status of the data holder. Fourth, the current rudimentary stage of the development and deployment of data governance mechanisms/models poses significant challenges and creates ambiguities about the legal form and administration of data spaces. Fifth, technical implementation of the rules and contracts relating to data access and sharing pose significant challenges. There exist no prior methodologies to efficiently incorporate norms into design choices or technical specifications.

1.2 Purpose and the structure of the document

This deliverable under Task 2.5 *Data Governance Requirements* focuses on data governance mechanisms/models and legal entitlements on data that affect data transactions.

Chapter II starts with the preliminaries first by elaborating on the concept of data governance and then by developing a conceptual framework consisting of *organisational*, *legal and regulatory,* and *technological* dimensions (Sections 2.1 and 2.2). The Chapter further provides a mapping of the existing and suggested forms of data governance models and mechanisms that have so far been introduced by the sectoral initiatives, legislative proposals and research projects (section 2.3). Based on the use cases provided in D2.9 of the Project, Section 2.4 explores how different types of data governance models and mechanisms (as collaborative data empowerment settings) could accommodate and support data access and sharing in the Transport Cloud. The analysis aims to shed light on the variables that have an impact on the data governance mechanisms and actions that are required to allow such data governance mechanisms to flourish in the Transport Cloud.

Chapter III focuses on legal frameworks which provide substantive rights and entitlements enabling trading, sharing and otherwise use of the data. These legal regimes, i.e., the personal data protection regime, IP rights (copyright and sui generis database right) and trade secret ("TS") protection, are the main legal frameworks that give control over data and thus enable data transactions.

Section 3.2 analyses how the legal regime for the protection of personal data**,** primarily the GDPR**,** could affect data transactions. When data could be linked to an identified or identifiable natural person, the rules on personal data protection limit the circulation and accessibility of data and thus affect both the control and usage as could be contemplated in a data transaction. Yet, the personal data protection regime may also serve as a facilitator for data transactions through data subject rights such as data portability, right to erasure, right to revocation of consent, and more generally the provisions of the GDPR on transparency and access. As these rights confer control over personal data, they help create a more trusted environment for data subjects who are willing to make their data available. Section 3.3 focuses on IP rights as a framework conferring exclusive rights on certain forms of information and thus providing the legal basis for a variety of data exchanges. Individual data items may fall under the definition of copyrightable work and more importantly, a database may qualify for protection under the *sui generis right* if the conditions set out by the Database Directive are satisfied. This part provides a comprehensive view of the applicability of the *sui generis* protection to machine-generated databases and highlights the relevant legal discussions. Section 3.4 analyses the TS Directive explaining how the legal protection of sensitive business information could act both as an enabler and also as an impediment to data transactions. While TS protection could foster data transactions by unlocking data that would be kept in the dark in the absence of TS protection, it is also possible that extensive reliance on TS could severely impede the efficient operation of data markets. Section 3.4 inquires under what conditions different types of data could satisfy the eligibility criteria for TS protection.

Chapter IV focuses on the DA proposal which is the second major legislative instrument of the emerging EU data governance regime following the DGA.[2] The DA proposal is horizontal legislation aiming to make data available to a wider range of stakeholders for innovative use. It introduces obligations to give access to data generated by the IoT devices, and lays out general rules regarding data contracts either entered into voluntarily or mandated by law. In relevant sections, alterations to the existing legal frameworks and the requirements relating to data transactions (contracts) introduced by the DA proposal are explained. Section 4.3 further inquires how this novel legislative initiative could affect the MobiDataLab project and the Transport Cloud.

[2] The Data Act proposal was released on 23 February 2022, in line with the 2020 European Data Strategy and the DGA, the proposal significantly reshapes the European regulatory framework for data.

2. Data governance framework in a legal perspective

2.1 What is data governance?

The proliferation of IoT devices, smart phones and online services has given rise to the collection of vast amounts of data that could be used to optimize various processes and enable new services in the mobility and transport sector. The exploitation of data has a fundamental impact in particular on how organizations operate and compete. While the flow of information is critical in such environments, in many cases the real-time and distributed nature of the system complicates the mechanisms and processes for protecting and controlling access to data.[3] This makes companies and institutions increasingly confronted with several strategic business requirements about how to regulate and govern data as an economic asset to maximize the value created.[4] Expanding data volumes from diverse sources compromises data quality and escalates risk exposure while increasing the pressure on organizations regarding compliance with the intense and complex regulatory landscape.[5]

In response, *data governance* emerges as a dedicated approach to coordinating and organizing both internal and external data-related activities.[6] Data governance could be defined as a system of

[3] Calo, Seraphin, Elisa Bertino, and Dinesh C. Verma, eds. *Policy-Based Autonomic Data Governance*. Lecture Notes in Computer Science 11550. Cham: Springer, 2019; Smichowski, Bruno. (2019). Alternative Data Governance Models: Moving Beyond One-Size-Fits-All Solutions. Intereconomics. 54. 222-227. 10.1007/s10272-019-0828-x

[4] Boris Otto, 'A Morphology of the Organisation of Data Governance ' [2011] ECIS 2011 Proceedings <https://aisel.aisnet.org/ecis2011/272>. Also see Zeljko Panian, 'Some Practical Experiences in Data Governance' [2010] World Academy of Science, Engineering and Technology 8. The term "governance", in general, refers to the way an organisation ensures that strategies are set, monitored, and achieved. Kristin Weber, Boris Otto and Hubert Österle, 'One Size Does Not Fit All---A Contingency Approach to Data Governance' (2009) 1 Journal of Data and Information Quality 4:1 The term [governanace] stresses a discontinuity from so-called "command-and-control" by the State, and acknowledges that a broader set of actors and institutions are involved.
Otto, B. 2011a. "Data Governance," *Business & Information Systems Engineering* (3:4), pp. 241–244.

[5] Benfeldt, O. (2020). *Polycentric governance of organizational data ventures: An organizing logic for data governance in the digital era*. Aalborg Universitetsforlag. Aalborg Universitet. Det Samfundsvidenskabelige Fakultet. Ph.D.-Serien.

[6]; Khatri, V., and Brown, C. V. 2010. "Designing Data Governance," *Communications ofthe ACM* (53:1), p. 148; Al-Ruithe, M., and Benkhelifa, E. 2017. "A Conceptual Framework for Cloud Data Governance-Driven Decision Making," in *Conference Proceedings - 2017 International Conference on the Frontiers and Advances in Data Science, FADS 2017* (Vol. 2018-January), pp. 1–6. ; rous, P., Janssen, M., and Vilminko-Heikkinen, R. 2016. "Coordinating Decision- Making in Data Management Activities: A Systematic Review of Data Governance Principles," in *International Conference on Electronic Government* (Vol. 9820), H. J. Scholl, O. Glassey, M. Janssen, B. Klievink, I. Lindgren, P. Parycek, E. Tambouris, M. A. Wimmer, T. Janowski, and D. Sá Soares (eds.), Cham: Springer International Publishing, pp. 115–125. ; Abraham, R., Schneider, J., and vom Brocke, J. 2019. "Data Governance: A Conceptual Framework, Structured Review, and Research Agenda," *International Journal of Information Management* (49), pp. 424–438.
https://vbn.aau.dk/ws/files/400433240/PHD_OB_E_pdf.pdf

Funded by the European Union

decision rights and *accountabilities* for information-related processes, executed according to *agreed-upon models*—describing who can take what actions with what information, and when, under what circumstances. As a set of policies and procedures, data governance aims to ensure that data is managed in a legally compliant, trustworthy and efficient way. The concept encapsulates the practice of organizing and implementing policies, procedures and standards for the effective use of an organization's structured and unstructured data assets.[7] Although there exists no consensual definition, the concept involves the design and implementation of rules and responsibilities, which specify how data as organizational assets may be treated. In a broader perspective, data governance refers to the exercise of authority and control over the management of data.[8]

Platform-based infrastructures and cloud-to-edge technologies significantly increase the potential of value creation from data— giving rise to a shift of focus to an inter-organizational perspective which includes novel forms of collaboration between various actors.[9] This is how the notion of data governance (as being closely linked to the external use of data, e.g., sharing between organizations) and the relevant technical infrastructures have been a topic of interest for entities that collect produce and rely on data for their activities.[10]

Data governance roots back in various traditions of IT governance, data quality management and information management. Differing from IT governance and data management, data governance is primarily concerned with value creation in a data ecosystem.[11] Hence, in addition to the internal management of data, data governance embraces the whole data lifecycle from collection to destruction (erasure). The primary focus is making data available (readily accessible) to a wider range of stakeholders in a form that complies with the regulatory framework (i.e., legislation, industry

[7] Dominik Lis and Boris Otto, 'Data Governance in Data Ecosystems – Insights from Organizations' [2020] AMCIS 2020 Proceedings <https://aisel.aisnet.org/amcis2020/strategic_uses_it/strategic_uses_it/12>.
[8] David Plotkin, *Data Stewardship: An Actionable Guide to Effective Data Management and Data Governance* (Second edition, Academic Press 2021); Rene Abraham, Johannes Schneider and Jan vom Brocke, 'Data Governance: A Conceptual Framework, Structured Review, and Research Agenda' (2019) 49 International Journal of Information Management 424; Majid Al-Ruithe, Elhadj Benkhelifa and Khawar Hameed, 'Data Governance Taxonomy: Cloud versus Non-Cloud' (2018) 10 Sustainability 95; Lis and Otto (n 7); Robert Seiner and an O'Reilly Media Company Safari, *Non-Invasive Data Governance* (2014) <https://www.safaribooksonline.com/complete/auth0oauth2/&state=/library/view//9781935504870/?ar> accessed 29 October 2021..
[9] Susha, Iryna; Gil-Garcia, J. Ramon. (2019). A Collaborative Governance Approach to Partnerships Addressing Public Problems with Private Data. Proceedings of the 52nd Hawaii International Conference on System Sciences 10.24251/HICSS.2019.350 Lis and Otto (n 7).
[10] Fabian de Prieëlle, Mark de Reuver and Jafar Rezaei, 'The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry' [2020] IEEE Transactions on Engineering Management 1.
[11] Meanwhile, data governance emerged as an organization-wide approach to data from various traditions of IT governance, data quality management and information management (Khatri and Brown 2010; Ladley 2012; Otto 2011a). Early conceptions of data governance focused on the implementation of formal rules and responsibilities, which specified decision-making and accountabilities within a series of decision domains regarding an organization's data assets (Khatri and Brown 2010). Much like IT governance (Weill and Ross 2004), data governance emphasized data as an organizational asset with the expressed objective of aligning data activities with business imperatives (Ladley 2012), making it well-suited for providing an organization-wide approach to data use. Scholars have even begun to highlight the potential of data governance for managing issues of privacy, data protection legislation and ethics (Abraham et al. 2019; Vydra and Klievink 2019).

standards and contractual rules).[12] Data governance approaches data as an organizational asset with the expressed objective of coordinating data for both internal and external activities with business imperatives.[13] The goals and objectives of data governance also include the development and deployment of the methods, set of responsibilities, and processes to standardize, integrate, protect, and store data.[14] As such data governance covers all participating groups, all types of *data* and *data use* and adapt to different business models and contexts.[15] In an overarching view, the concept brings together the efforts which enhance legal compliance and protection of business interests while achieving the widest possible data fluidity.[16]

So far, the research on data governance mainly remains conceptual mostly focusing on how to design top-down programs for organisational data use.[17] The current perspective is generally narrowed down to normative aspects such as decision-making rights and formal roles. This proves ill-equipped to cope with the turbulent, multifaceted reality of organizational data use. Critical of these unilateral and normative approaches, some scholars indicate that hierarchical structures alone may be inadequate for data resource management in the information age.[18] This line of scholarship voices concerns about the growing datafication which needs rigorous questioning and draws attention to the socio-political, cultural, economic and ethical implications following increasing data exploitation.[19] A budding line of research, known as *critical data studies*, investigates data-centric technologies and their pervasive infrastructures as 'assemblages' with agency that inflect and interact with society and individuals. [20] This interdisciplinary approach put special emphasis on the power asymmetries among the stakeholders and the underlying rationales and system of thoughts.[21] These power asymmetries in data governance are of significant concern as they are the source of unfair and exploitative practices.[22] In line with this, there is also growing emphasis on the notion of *data justice*

---

[12] Abraham, Schneider and vom Brocke (n 8).
[13] Benfeldt, O. (2020). *Polycentric governance of organizational data ventures: An organizing logic for data governance in the digital era*. (Ladley 2012),
[14] Thor Olavsrud, 'Data Governance: A Best Practices Framework for Managing Data Assets' (*CIO*, 18 March 2021) <https://www.cio.com/article/3521011/what-is-data-governance-a-best-practices-framework-for-managing-data-assets.html> accessed 29 October 2021.
[15] Sung Une Lee, L Zhu and R Jeffery, 'Data Governance Decisions for Platform Ecosystems', *HICSS 2019 : Proceedings of the 52nd Hawaii International Conference on System Sciences* ([The Conference] 2019).
[16] Weber, Otto and Österle (n 4).
[17] Brous et al. 2016; Otto 2011d, 2011c; Weber et al. 2009 Alhassan, I., Sammon, D., and Daly, M. 2019. "Critical Success Factors for Data Governance: A Theory Building Approach," *Information Systems Management* (36:2), pp. 98–110. ; Mikalef, Pappas, et al. 2020
[18] Levitin and Redman (1998), Begg, C., and Caira, T. 2011. "Data Governance in Practice:: The SME Quandary Reflections on the Reality of Data Governance in the Small to Medium Enterprise (SME) Sector," in *The European Conference on Information Systems Management*, , September, pp. 75–VIII ; Buffenoir and Bourdon 2013
[19] Benfeldt, O. (2020). *Polycentric governance of organizational data ventures: An organizing logic for data governance in the digital era*.
[20] (Iliadis and Russo 2016).
[21] Marina Micheli, Marisa Ponti, Max Craglia, Anna Berti Suman, 'Emerging Models of Data Governance in the Age of Datafication' (2020) 7 Big Data & Society 2053951720948087.
[22] Alison Holt (ed), *Data Governance: Governing Data for Sustainable Business* (2021); Lorena Elena Stanescu and Raluca Onufreiciuc, 'Some Reflections on "Datafication": Data Governance and Legal Challenges' (2020) 7 European Journal of Law and Public Administration 100; For more on data management and governance, see Al-Ruithe, Benkhelifa and Hameed (n [8]); Another relevant concept is data stewardship which refers to the operational aspect of Data Governance— and formalizes accountability for managing

stemming from pre-existing sociocultural biases which increasingly become embedded in certain data-centric technologies.[23] The public administration and law enforcement practices where citizens are managed as risk-scores are of special attention as they undermine procedural legal safeguards.[24] In sum, there is growing interdisciplinary research under the banner of *critical data studies* which view data-centric technologies as techno-social assemblages claiming human experience as the raw material for their hidden practices of knowing, controlling, modifying and commodifying human behaviour.[25]

Despite the increasing attraction of the concept, organizing data use involves various imperatives which are complex and at times even contradictory as the stakeholders have diverse interests. Primarily, there is a need for a consensual definition and a general framework. Moreover, although it is generally agreed that data governance is both promising and necessary in the digital era, there exist significant ambiguities and knowledge gaps about how data governance unfolds in practice. Due to the special emphasis given to the concept by the EU Commission, it is expected that the prospective research on data governance will explore the broader aspects of data governance.[26] As data governance emerges as one of the core concepts in the 2020 European Data Strategy, there are already significant efforts that emphasize the potential of data governance for managing complex issues of privacy, personal data protection, and data sharing in organizational data use.[27]

As knowledge on organizational data use is scattered across diverse research streams, there is still a need for more original theorising for an overarching treatment of data governance both as a concept and as a set of practices. This will enhance cross-fertilization and consolidation of empirical findings across contexts.[28] In addition to enduring theoretical and conceptual issues, empirical investigations show that many organisations have second thoughts in terms of the complexity of the

information resources on behalf of others and for the best interests of the organization. Rupa Mahanti, *Data Governance and Data Management: Contextualizing Data Governance Drivers, Technologies, and Tools.* (Springer 2021).

[23] Berry, D. M. 2019. "Against Infrasomitization: Towards a Critical Theory of Algorithms," in *Data Politics: Worlds, Subjects, Rights*, D. Bigo, E. F. Isin, and E. Ruppert (eds.), London: Routledge, pp. 43–63; Crawford, K., Miltner, K., and Gray, M. L. 2014. "Critiquing Big Data: Politics, Ethics, Epistemology," *International Journal of Communication* (8), pp. 1663–1672

[24]Dencik, L., Hintz, A., and Cable, J. 2016. "Towards Data Justice? The Ambiguity of Anti-Surveillance Resistance in Political Activism," *Big Data & Society* (3:2), pp. 1–12.; Taylor, L. 2017. "What Is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally," *Big Data & Society* (4:2), p. 205395171773633. For thee thical dimension see Zigon, J. 2019. "Can Machines Be Ethical? On the Necessity of Relational Ethics and Empathic Attunement for Data-Centric Technologies," *Social Research: An International Quarterly* (86:4), Johns Hopkins University Press, pp. 1001–1022. ; Knox, H., and Nafus, D. 2018. "Introduction: Ethnography for a Data-Saturated World," in *Ethnography for a Data Saturated World*, p. 30.

[25] (Taylor 2017; Zuboff, S. 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization," *Journal of Information Technology* (30:1), pp. 75–89; Zuboff, S. 2018. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, (First edition.), New York: PublicAffairs

[26] Benfeldt, O. (2020). *Polycentric governance of organizational data ventures: An organizing logic for data governance in the digital era;* Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., and Janowski, T. 2020. "Data Governance: Organizing Data for Trustworthy Artificial Intelligence," *Government Information Quarterly*, Elsevier, p. 101493.

[27] Abraham et al. 2019; Addis and Kutar 2018; Vydra and Klievink 2019).

[28] Benfeldt, O. (2020). *Polycentric governance 9 . Also see* Grover, V., and Lyytinen, K. 2015. "New State of Play in Information Systems Research: The Push to the Edges.," *Mis Quarterly* (39:2), pp. 271–296.

applicable laws and adhering to certain data standards.[29] Relevant and timely guidance on how to organize data-related activities and manage competing interests on data is a pressing need.

2.2 The conceptual framework of data governance

Based on the contributions of the disciplines of Science and Technology Studies ("STS"), organisational design, management studies, information science and IT law in general, the conceptual framework of data governance may be studied in three dimensions: i) *organisational*, ii) *legal and regulatory,* and iii) *technological*.[30]

2.2.1 Organisational dimension

The organisational dimension in data governance relates to a wide range of formal, functional and administrative goals—covering *decision-making and control*, *organisational form*, and *roles and responsibilities*.[31] *Decision-making and control* focus on the hierarchies and controls embedded in the general management architecture. The organisational form has both legal and *administrative/managerial aspects*. When organised as a separate entity, the organisational form could be a *company, association, public enterprise*, or other *data intermediaries sharing* as defined in the DGA proposal.[32] Data governance systems could also be classified as *centralised, decentralised, federated,* or *project-based (ad-hoc)*.[33] Various types of *models and mechanisms* are being enacted, discussed and experimented with to address the problems relating to the *form* and *control* in data governance settings. The roles and responsibilities in data governance relate to a network of stakeholders which includes a wide spectrum of organisations from industry players of various sizes and associations to standardisation bodies, regulators, and digital innovation hubs.

Under the emerging EU data governance regime, the organisational dimension of data governance is structured around the concept of "data space"[34]. Aiming to create an interoperable data sharing

[29] Vilminko-Heikkinen, R., and Pekkola, S. 2019. "Changes in Roles, Responsibilities and Ownership in Organizing Master Data Management," *International Journal of Information Management* (47), pp. 76–87.
[30] This part is an adaptation and elaboration of the framework developed in D3.8 (also by KU Leuven) of the EUH4D Project H2020.
[31] Otto (n 4); Al-Ruithe, Benkhelifa and Hameed (n 8).
[32] European Commission, 'Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European data governance (Data Governance Act)' COM/2020/767 final, ("Data Governance Act"). The DGA proposal constitutes the first legislative initiative under the European Data Strategy, introducing a broad framework of horizontal, cross-sectoral measures, relevant for all data spaces while leaving room for vertical, sector-specific rules. For further information, see Deliverable D2.1 "Legal and regulatory data sharing gap analysis" and Julie Baloup, Emre Bayamlıoğlu, Aliki Benmayor, Charlotte Ducuing, Lidia Dutkiewicz, Teodora Lalova, Yuliya Miadzvetskaya, Bert Peeters, 'White Paper on the Data Governance Act' (CiTiP KU Leuven 2021) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3872703,> accessed 7 April 2022.
[33] Otto (n 4).
[34] European Commission, Commission Staff Working Document on Common European Data Spaces, Brussels, 23.2.2022 SWD(2022) 45 final.

Funded by the European Union

environment that enables data reuse across sectors, the European Data Strategy has its centrepiece the concept of *data spaces* as federated (decentralized) data ecosystems with shared legal, operational, functional agreements and technical standards.[35] The common European data spaces in each relevant sector or domain provide a set of interoperable data-sharing applications either by internal development by the stakeholders or through a certified software vendor, data broker, or marketplace. Data spaces enable interactions between multiple actors of the data ecosystem and in particular connect data providers with data users, serving as the building blocks upon which an array of firms and entities develop complementary products, technologies or services.[36] The common European mobility data space may be seen as the umbrella structure where the Project output will be implemented and put into practice. It will facilitate access, pooling and sharing of data from existing and future transport and mobility databases.[37]

2.2.2 The regulatory and legal dimension

The regulatory and legal dimension covers a plethora of norm-setting/providing instruments (both domestic and EU level), ranging from administrative decrees/orders, court decisions, contractual commitments, intra-institutional rules. Among those, the most significant EU legislation are the General Data Protection Regulation, the e-Privacy Directive; rules on competition law; Public Sector Information Directives (including the 2019 Open Data and Public Sector Information Directive); Regulation on the free flow of non-personal data; Legislation concerning digital platforms and/or intermediaries (e-Commerce Directive, Platform-to-Business Regulation and the recent Digital Services Act Package). The proposals for a DGA[38] and the DA (which will be analysed in section 4.2) are the most prominent regulatory instruments that make up the essential pillars of the emerging EU data governance regime.

2.2.3 The technological dimension

[35] Lars Nagel and Douwe Lycklama, 'Design Principles for Data Spaces - Position Paper' (Zenodo 2021) <https://zenodo.org/record/5105744> accessed 15 November 202, 93.Alexandru Ioan Cuza University of Iasi, Lorena Elena Stanescu, Raluca Onufreiciuc, and University of Bucharest. 'Some Reflections on "Datafication": Data Governance and Legal Challenges'. *European Journal of Law and Public Administration* 7, no. 1 (31 July 2020): 100–115. https://doi.org/10.18662/eljpa/7.1/118.Inge Graef, Martin Husovec and Jasper van den Boom, 'Spill-Overs in Data Governance: Uncovering the Uneasy Relationship between the GDPR's Right to Data Portability and EU Sector-Specific Data Access Regimes' (2020) 9 Journal of European Consumer and Market Law 3.

[36] Lee, Zhu and Jeffery (n 15).

[37] See below Section 4.1.

[38] Analysed in D2.1 of the MobiDataLab project, "Legal and regulatory data sharing gap analysis".

The technological dimension covers a wide range of elements related to the operationalisation of data governance. The software architecture and the design choices are the essential tools to implement legal and contractual rules on data governance. These rules relate to various technical issues such as availability, reliability, security, privacy, quality, compatibility, portability, control, auditing and integrity of data.

As to the technical solutions, *cloud-to-edge* architectures address the increasing demand for speed and bandwidth in data governance. Considering the latency and bandwidth limitations inherent in transmitting data up to a cloud service, edge computing systems and Internet of Things (IoT) devices bring about a swing back to decentralized models. The need for real-time monitoring and analytics capabilities for certain services requires a delicate balance between edge-based devices and cloud systems. The 2020 EU Data Strategy[39] and the Member States' Joint Declaration on Cloud[40] acknowledge the strategic role of cloud-to-edge technologies in the successful digital transformation. There is a special emphasis on the deployment of trusted, interoperable and sustainable cloud-to-edge capabilities compliant with EU rules. As low-power computing solutions, edge technologies are found to be important for the twin transitions to a green and digital Europe.

On the technical dimension, interoperability and standardization are repeatedly emphasized as the objectives that should be pursued throughout the entire lifecycle of data. Data governance systems are required to implement modules for auditing, reporting and logging. The design architecture should align with the defined permissions and restrictions in the applicable contracts, terms of use and data policies.

2.3 Data governance models and mechanisms

Models and mechanisms which bring together technical, organisational and normative(regulatory) elements of data governance are essential for the provision, access and sharing of data in a collaborative and trustful environment—encompassing all stakeholders and addressing matters related to quality control, and the possible use cases of data.[41]

As an overarching concept, governance models and mechanisms include the *formal structures* integrating data management functions, *processes and procedures* for decision-making and monitoring, and *practices* supporting the active participation of and collaboration among the stakeholders.[42] Data governance models and mechanisms mainly aim to address challenges, i.e., control over data at *intra-* and *inter-organisational* level; decision-making within the governance structure; interoperability and traceability of data; data quality; value creation, and legal

[39] European Commission, 'A European Strategy for Data' 19 February 2020, COM(2020) 66 Final'; European Commission, 'European Data Strategy' <https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en> accessed 07 April 2022.
[40] European Commission, 'Towards a next Generation Cloud for Europe | Shaping Europe's Digital Future' <https://digital-strategy.ec.europa.eu/en/news/towards-next-generation-cloud-europe> accessed 07 April 2022.
[41] de Prieëlle, de Reuver and Rezaei (n 10). See also D3.4 "Data sharing business and revenue models" and the business and legal structures described therein
[42] Abraham, Schneider and vom Brocke (n 8).

Funded by the European Union

compliance.[43] Given that there exists no consensual definition or optimal solutions, at the practical level we see a combination of characteristics of various models and mechanisms depending on the specificities of the particular data use.

2.3.1 Data pools

Data pooling could be defined as a "data-sharing system which involves an element of reciprocity, whereby at least some companies contribute data". [44] It is a system or setting where two or more entities undertake to transfer and aggregate data in a jointly controlled medium. The term also refers to a form of industrial data sharing where entities share their informational resources within a data e-ecosystem.[45]

Data pools or similar concepts are implicitly referred to in various reports, policy documents and legislative texts of the European Commission as a collaborative and strategic approach where participants create a secure and exclusive environment to exchange data.[46]

Data pooling refers to a communal approach to data sharing aiming to create a model under which data subjects and data holders jointly decide on norms and principles on how their data shall be collected and used and those who will access this communal source. As such, they do not generally contain mechanisms for commercial exploitation or monetisation. Data pools may be viewed akin to 'patent pools' where firms agree to license complementary patents through a single agreement and at a standard royalty fee.[47] Through data pooling, stakeholders can have access to a scale of data that would not have been possible without the aggregation in a common *reservoir*. [48] Data pooling

---

[43] Sayogo, Djoko and J. Ramon Gil-Garcia. (2015). Analyzing the Influence of Governance Structure Determinants on the Success of Inter-Organizational Information Sharing Initiatives. Paper presented at the 48th Hawaii International Conference on System Sciences (HIICS), organized by the College of Business, University of Hawai'i at Mānoa, Big Island, Hawaii, USA, January 5-8.

[44] European Commission, Directorate General for Communications Networks, Content and Technology and others, 'Study on Data Sharing between Companies in Europe: Final Report' (Publications Office of the EU 2018) <https://data.europa.eu/doi/10.2759/354943> accessed 7 April 2022.

[45] The term is also used to refer to a form of industrial data sharing where "firms agree to share their digitalized information in reference to a given service or generally in an industry, or within an e-ecosystem". Bjorn Lundqvist, 'Competition and Data Pools' (2018) 7 Journal of European Consumer and Market Law 146.

[46] Data pooling is specifically mentioned in the *Explanatory Memorandum* of the proposed AI Act—acknowledging that the trusted mechanisms and services for the reuse, sharing and *pooling of data* are essential for the development of data-driven AI models of high quality. European Commission, 'Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts COM/2021/206 Final ("AI Act")'.

[47] Oscar Borgogno and Giuseppe Colangelo, 'Data Sharing and Interoperability: Fostering Innovation and Competition through APIs' (2019) 35 Computer Law & Security Review 12 <https://www.sciencedirect.com/science/article/pii/S0267364918304503> accessed 7 April 2022.

[48] Margrethe Vestager, 'Big Data and Competition' (Speech at EDPS-BEUC Conference on Big Data, Brussels, 29 September 2016.<https://web.archive.org/web/20210314065315/https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/big-data-and-competition_en> accessed7 April 2022.

enables SMEs and start-ups to tap into vital resources without incurring high sunken costs Hence data pools are generally regarded to have pro-competitive effects as they make (otherwise undisclosed) data available to a wide span of market players. [49] There exist various experimental efforts establishing organisational structures where data can be aggregated, shared and managed by the peers themselves, based on community norms and bylaws set up by those who contribute to the data pool.[50]

It should be noted that despite being a promising a useful model at the conceptual level, the exact function of data pools and their legal status is not clear.[51] Some scholars argue that data pools should be distinguished from the broader concept of *trusted intermediaries or* "data-sharing platforms" which also engage in data transformation or anonymisation.[52] Although such distinctions may be useful at the conceptual level, in practice, data pools may be organised in numerous ways combining features from various data governance mechanisms and business models. [53] In sum, there is no clarity or agreement as to the legal status or concrete mechanisms governing such pooled resources. A plethora of terms and concepts (with or without explicit mentioning) may be associated with the models or mechanisms where actors collaboratively pool their data.

2.3.2 Data commons

The concept, data commons, refers to management-related characteristics of *joint* or *collaborative governance mechanisms* between participants, where many parameters such as the technical level of aggregation and interoperability, the conditions for access and use of data are of concern.[54] It is a communal approach for the collective governance of the shared sources which could be combined with other models and organised both in ways centralised or decentralised. [55] As such they present significant similarities with data pools as collaborative models.

Commons-based data governance models and mechanisms are efficient in curbing the anti-competitive conduct exercised by the strong actors in the data ecosystem. Successful

[49] J. Crémer, Y.-A. de Montjoye, and H. Schweitzer, "Competition policy for the digital era," Publications Office of the EU, 2019. Accessed: 7 April 2022. [Online]. Available: https://data.europa.eu/doi/10.2763/407537.
[50] Chih-Hsing Ho Tyng-Ruey Chuang, Governance Of Communal Data Sharing, In Angela Daly, S. Kate Devitt and Monique Man(eds.), *Good Data*, the Institute of Network Cultures, Amsterdam, 2019.
[51] Michael Mattioli, 'The Data-Pooling Problem' (2017) 32 Berkeley Technology Law Journal 179; Borgogno and Colangelo (n 48).
[52] Also referred to as data sharing services in the DGA Proposal. See European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act)' COM/2020/767 Final ("DGA Proposal").
[53] Data pools could both be centrally administered or may be arranged similar to *data commons* without need for a central authority as will be explained in the following sub-section below. Borgogno and Colangelo (n 48). Also see Heiko Richter and Peter R Slowinski, 'The Data Sharing Economy: On the Emergence of New Intermediaries' (2019) 50 IIC - International Review of Intellectual Property and Competition Law 4
[54] Jennifer Shkabatur, 'The Global Commons of Data' (2019) 22 Stanford Technology Law Review 354.
[55] Elinor Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge University Press 1990); Jennifer Shkabatur (n 55). Michael J Madison, Brett M Frischmann and Katherine J Strandburg, 'Constructing Commons in the Cultural Environment' [2010] CORNELL LAW REVIEW 657.

implementations of data commons exist in the automotive sector where vehicle manufacturers generally tend to foreclose access to in-vehicle data.[56] A *neutral server* (a.k.a *shared server)* model based on the "extended vehicle concept"[57] has been put to use for the collaborative management of stakeholders such as spare parts sellers or repair service providers or drivers.[58] Data commons could also be administered by an independent operator to act as an intermediary among the stakeholders.[59] Commons-based communal approaches are seen as a solution for the power asymmetry between the stakeholders in a data eco-system with unequal means of access and control over data.[60]

Where commons-based models are in use—rather than mere passive data contributors— smaller actors (e.g., individuals, SMEs and start-ups) become active participants and decision-makers in the data value chain. Lastly, it should be noted "open data" as laid out by the Open Data and PSI Directive does not equal to *data commons* since the Directive has no management or control dimension and it focuses on certain permitted types of access.[61]

2.3.3 Data trusts

The concept of *data tru*st emerges as a form of data stewardship based on the Common Law institution which binds the *trustee* and the *trustor* under certain fiduciary duties.[62] In this legal structure, the *trustee* owes duties to act in the best interests of the *trustors*. That said, the general reference to "data trusts" both in the literature and in the EU policy documents is not necessarily limited to the rigid structure of Common Law but rather covers the provision of a service under an

[56] Considering this monopolistic tendency of manufacturers, the recently proposed Data Act provides a right of Access to the data generated by IoT devices. Yet the right is confined to device users and does not directly include other stakeholders. See below section 4.2.1.

[57] The 'extended vehicle concept' provides a standardisation for access to car data by third-parties. It is filed as an ISO standard (20078–1.) enabling different services and interfaces to access to both anonymous and personalised vehicle data irrespective of the system they are using.

[58] M McCarthy, M Seidl, S Mohan, J Hopkin, A Stevens, F Ognissanto, 'Access to In-Vehicle Data and Resources' (2017) Publications Office of the European Union.

[59] Similar commons-based approaches have been proposed for electricity data management. Industry examples underline the regulatory role of the technology owing to the lack of a general access regime for data.

[60] A successful example of data commons come from the agricultural sector where farmers do not have any control or rights on the data generated throughout their activities while service providers and/or device manufacturers acquire intellectual property rights as they collect and aggregate farmers' data Jeremiah Baarbé, Meghan B and Jeremy de Beer, 'A Data Commons for Food Security' (Social Science Research Network 2017) Proceedings of the 2017 IASC Conference ; Open AIR Working Paper No 7/17 ID 3008736.

[61] However the OECD report refers to "open data" as a typical example of data commons. OECD, 'Enhancing Access to and Sharing of Data : Reconciling Risks and Benefits for Data Re-Use across Societies' (OECD Library 2019) <../sti-2019-1215-en/index.html>.

[62] See Kieron O'Hara, 'Data Trusts Ethics, Architecture and Governance for Trustworthy Data Stewardship' (University of Southampton 2019) 1; Element AI and NESTA, 'Data Trusts A New Tool for Data Governance' (2019) <https://hello.elementai.com/rs/024-OAQ-547/images/Data_Trusts_EN_201914.pdf>. An example of the data trust is UK Biobank which holds data from about 0.5m people and it includes the number of 946 researchers using its data in its annual accounts of 2018.

organisational structure as a type of *independent stewardship* of data.[63]Data trusts integrate data from multiple sources, involve various kinds of stakeholders, and could employ both collaborative and centralised models depending on the context and purpose.[64]

As a special form, '*public'* or '*civic'* trust implies the establishment of a relationship between data subjects or legal entities and public bodies. In this model of data governance, a public body accesses, aggregates and uses data held by natural persons and legal entities. Trustors are assured that public actors are capable of keeping their personal data or confidential business information safe and secure.[65] The concept of *public data trusts*, where a public authority acts as the steward, resembles the commons-based models as they both connote some degree of legal and technical administration of data.

2.3.4 Data marketplaces

In general, markets help mitigate the information asymmetries among the actors, as they provide a mechanism to inform potential buyers about the quality, scope and content of the data. [66] Despite the lack of a consensual definition, *data marketplace* refers to contractual and technical settings where data users and providers meet and enter into transactions on free terms and equal basis. A data marketplace may be viewed as a "match maker" enabling sellers to offer data products and services, and buyers to find and acquire data.

Data marketplaces usually come into being as electronic venues or platforms which provide the infrastructure that allows participants to meet and define their terms of data use and other essential elements of a data contract.[67] Data marketplaces reduce transaction costs by automating exchanges and thus help bypass the costly rights clearance process. [68]

---

[63]Jack Hardinges, Peter Wells, Alex Blandford, Jeni Tennison, Anna Scott, '"Data Trusts: Lessons From Three Pilots", Open Data Institute' (2019) <https://docs.google.com/document/d/118RqyUAWP3WIyyCO4iLUT3oOobnYJGibEhspr2v87jg/edit> accessed 7 April 2022; Also see Sylvie Delacroix and Neil D Lawrence, 'Bottom-up Data Trusts: Disturbing the "One Size Fits All" Approach to Data Governance' International Data Privacy Law <http://academic.oup.com/idpl/advance-article/doi/10.1093/idpl/ipz014/5579842>.

[64] Element AI and NESTA (n 62) 9.

[65] Massimo Craglia, Henk J Scholten, Marina Micheli, Jiri Hradec, Igor Calzada, Steven Luitjens, Marisa Ponti, Jaap Boter, *Digitranscope the Governance of Digitally-Transformed Society* (Publications Office of the EU 2021) 47 <https://doi.org/10.2760/503546> accessed 7 April 2022; Also see Delacroix and Lawrence (n 63).

[66] Richter and Slowinski (n 54) 13.

[67] Johannes Deichmann, Kersten Heineke, Thomas Reinbacher, Dominik Wee, 'Creating a Successful Internet of Things Data Marketplace' (*McKinsey Digital*, 7 October 2016) <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/creating-a-successful-internet-of-things-data-marketplace> accessed 7 April 2022.

[68] Florian Stahl, Fabian Schomm, Gottfried Vossen, Lara Vomfell , 'A Classification Framework for Data Marketplaces' (2016) 3 Vietnam Journal of Computer Science 137, 137; OECD (n 61); European Commission, 'SMART 2019/0024 | D2, Impact Assessment on Enhancing the Use of Data in Europe, Report

The concept of a data marketplace, as a technical setting, may also be deployed to monitor participants, supervise contracts and enforce rights and constraints as prescribed by law or in a specific contract. Data marketplaces are operationalised through a technical (software) architecture implementing the transactional model.[69] Different kinds of technical solutions and services could be offered through data marketplaces—enabling transparent tracking of all data-related transactions. [70] As a consequence, a data marketplace could also be an integral part of other governance models, especially where data is exchanged within a commercial or business context.

2.3.5 Data cooperatives

In November 2020, the European Commission adopted the DGA proposal [71] as its first legislative initiative under the European Data Strategy. The proposal introduces a broad framework of horizontal, cross-sectoral measures, relevant to all data spaces while leaving room for vertical, sector-specific rules.[72] The proposal defines a certain type of data intermediation activities as *data sharing services* (DSS) and provides for a compliance and notification regime for entities falling under this category. Article 9 of the Proposal introduces two types of *data sharing service providers* namely "data intermediaries" and "data cooperatives".[73]

Data cooperatives are of particular importance as a collaborative form of data governance to enable access to data in a legally compliant way. Data cooperatives support their members by guiding them to strengthen their negotiating position before consenting or agreeing to data processing.[74] According to the proposal, data cooperatives facilitate the pooling of data by individuals, one-person companies or micro, small and medium-sized enterprises for their mutual benefit.[75] Data cooperatives are entrusted with the duty to provide know-how on data sharing to small businesses and establish mechanisms that would represent and protect the interests of their members.

The Proposal is criticised for its ambiguities concerning the legal form and organisational structure of data cooperatives. There is no clarity on whether the Proposal refers to *cooperative societies* as

---

on Task 1 – Data Governance' (2020) <https://www.asktheeu.org/en/request/9101/response/30449/attach/5/ANNEX%20I.pdf> accessed 7 April 2022 .

[69] Ibid.

[70] Nagel and Lycklama (n 35); Michael J Burstein, 'Exchanging Information Without Intellectual Property' 91 Texas Law Review 56.

[71] European Commission, DGA Proposal (n 53).

[72] European Commission, 'A European Strategy for Data'(n 39).

[73] DGA Proposal, Art.9

[74] Some examples of already existing entities under the name data cooperative are: MiData 'Home' (*MIDATA*) <https://www.midata.coop/en/home/> accessed 7 April 2022; 'SalusCoop' (*SalusCoop*) <https://www.saluscoop.org> accessed 7 April 2022; Niels-HHDC, 'HHDC - Holland Health Data Coöperatie - Beheer de sleutel tot jouw data' (*HHDC*) <http://hhdc.nl/> accessed 29 October 2021; The Good Data Cooperative 'TheGoodData | Enjoy Your Data' (*TheGoodData*) <https://www.thegooddata.org/> accessed 7 April 2022.

[75] European Commission, 'A European Strategy for Data' (n 39) 10.

Funded by the European Union

established by the Statute for a European Cooperative Society (ECS) as a form of business entity.[76] It is not clear from the proposal what is exactly meant by "data cooperative", and it is not possible to tell whether the legislature contemplates an entity directly within the ECS framework or whether the term is used in a rather generic sense.[77]

2.3.6 Data altruism organisations

Another novel type of entity introduced by the DGA proposal is the *data altruism organization*. As a means to foster access to data for the public good, the Proposal defines *data altruism* as *the consent by data subjects* to process their personal data, or the *permissions of other data holders* to allow the use of their non-personal data for purposes of general interest without reward.[78] The definition covers both personal and non-personal data while the respective provisions generally focus on or imply personal data. This gives rise to questions about how non-personal data would be used for altruistic purposes.[79] Regarding the use of personal data, data altruism organisations are required to comply with GDPR, meaning that data transactions for altruistic purposes are also subject to revocation of consent.[80] Hence, consent/permission management is an essential task in data altruism organisations. A further issue comes with the purpose limitation principle which requires specifics as to the purposes of processing and broad or open-ended definitions are generally not found to be compliant. Hence the question arises whether the purpose of altruistic ends may justify a more flexible approach to permit broad consent.[81]

Data altruism organisations are established to pursue goals that are of "general interest" such as protection of the environment or improving mobility. This gives rise to the question of whether and how the *general interest* in the DGA differs from the notion of *public interest* in Art. 6(1)(e) of the GDPR which serves as a ground for personal data processing without consent. It could be surmised that the concept of *general interest* needs to be interpreted to cover a wider range of activities in comparison to the *public interest*. Otherwise, the provision in the DGA would be redundant (at least so far as personal data is concerned) since the GDPR already permits personal data processing for public interest without consent. [82] Adding more to the existing ambiguity, the Impact Assessment on the DGA uses the term "*public* interest". Considering that the activities serving general interest do

[76] Council Regulation (EC) N° 1435/2003 of 22 July 2003 on the Statute for a European Cooperative Society (SCE), [2003] OJ L 207/1. The European Cooperative Society (ECS) is offered as a legal form to reduce existing cross-border obstacles for cooperatives and facilitate operation across Europe. unlike companies, members of ECSs do not control the organisation according to their capital contribution but the management is based on the principle of "one member, one vote"

[77] Baloup and others (n 32) 43. accessed 7 April 2022.

[78] DGA proposal, Art. 2(10).

[79] Rec. 38 specifies that "Data altruism organisations [...] should be able to collect relevant data directly from natural and legal persons or to process data collected by others".

[80] See above 4.2.1.

[81] See GDPR Rec. 33, DGA proposal Rec. 36, DGA Art.22 and European Data Protection Board, 'Guidelines 05/2020 on Consent under Regulation 2016/679 (Version 1.1)' (2020) <https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf> accessed 7 April 2022 ; For more on this, see Baloup and others (n 32) 37–47.

[82] SWD(2020) 295 final, 72.

Funded by the European Union

not necessarily exclude commercial or for-profit models, the provision in its current form may be practically difficult to implement without further clarification.[83]

As to the legal form, similar to data cooperatives and other DSS under the DGA proposal, data altruism organisations are also required to be established as independent legal entities.[84] Regarding this requirement, the ambiguities as to the exact scope of the activities that need to be handled by the independent entity may create challenges at the implementation stage.

[83] Baloup and others (n 32) 43. The DGA proposal has been explained in more detail in Deliverable D2.1 "Legal and regulatory data sharing gap analysis".
[84] DGA proposal, Art. 16(c).

Funded by the European Union

2.4 Data Governance models and Transport Cloud use cases

2.4.1 The Transport Cloud in general

MobiDataLab aims to provide mobility stakeholders with a methodology together with sustainable tools that foster a business and social culture of data sharing and access. Within this approach, the MobiDataLab Transport Cloud is a cloud-based prototype platform for making available transport and mobility data. In a federated approach, the developed tools and methodologies aim to balance long-term strategic shortcomings while also addressing the immediate requirements of the Project. [85]

The technical implementation of the Project, Transport Cloud, is designed to develop models which reduce and, where possible, remove current technical limitations in a legally compliant framework.[86] The design principles and the technical guidelines allow the platform's functionalities to be aligned with the requirements of the stakeholders represented in the Project (i.e., authorities, operators, MaaS companies and developers). The Transport Cloud addresses various challenges of data governance by using techniques of data fusion, anonymisation/ pseudonymisation, geographical and semantic enrichment, and standardisation. The overall aim is to improve the quality, accessibility and usability of mobility data. MobiDataLab project is built upon the fundamental principle of data distribution—offering a single data entry point for a multitude of data sources, with four main actors in its generic form namely, *administrator, developer, data consumer,* and *data/service provider.*

As a major requirement of the Project MobiDataLab, the Transport Cloud is designed to adopt an agnostic approach meaning that any suitable cloud platform may be used to support its implementation. This agnostic approach comes with certain challenges in that the system architecture needs to address diverse use cases and actors with diverging interests together with specific data needs that might be hard to reconcile. To address the legal issues relating to the use of personal data, the Transport Cloud must ensure that personal or sensitive data accessed or processed by the platform is only made available within the limits of the consent given by the data subject. The Project implements data anonymisation mechanisms and modules which could be deployed by the platform stakeholders when there is a need for access to or processing personal data. Regarding anonymisation, it is important to note that anonymisation techniques have certain drawbacks and trade-offs. Given enough time and resources, it could be possible to deanonymize data. In addition, based on the techniques used, the anonymisation process may diminish the informative character of data or data may cease to be useful for a specific purpose. There is always a need for a sense of proportionality in deciding which anonymisation techniques will be used and to what extent.

[85] D4.1 Transport Cloud Architecture Dossier V1.

[86] The Transport Cloud illustrates the design principles and technical guidelines that allow the platform's functionalities to be aligned with the requirements of the stakeholders represented in the project (i.e., authorities, operators, MaaS companies and developers). See "D2.9: Use cases definition (v1)" and "D2.6 : Report on enabling technologies for Transport Cloud", covering the basic functional and non-functional expectations from the platform in order for the report to conform to the requirements of those use cases.

Serialisation and standardisation are further essential tenets that underpin the design of the Transport Cloud, providing benefits in terms of costs, availability of data, and open documentation. The Project especially selects well-established open-source tools, technologies, and processes that align with the objectives of serialisation and standardisation**.**

As a cloud implementation strategy, the MobiDataLab project focuses on a set of common services that can be used to implement the functionalities of the Transport Cloud. These services provide the standard information technology tools required to build the Transport Cloud infrastructure, regardless of whether the infrastructure will be hosted on-premise, virtualised on some cloud platform, or some hybrid between these two approaches. As an agnostic approach, Transport Cloud's functionalities may be implemented regardless of the approach chosen to implement the underlying infrastructure.

Below, we provide an assessment of the *data governance models and mechanisms* by analysing their suitability for the Transport Cloud *use cases*. The use cases taken from D2.9 of the Project are standard examples demonstrating different functionalities of the Transport Cloud and thus, provide a concrete testbed to evaluate the applicability of data governance models.

2.4.2 Use case 1: Optimisation of Transport flow and ETA

Under this scenario, the data is used for computing Estimated Time of Arrival (ETA) which is highly relevant in optimising, monitoring, and managing transport flow. Package delivery and passenger carrier services use the system to predict the ETA for a package, vehicle or passenger at a given destination. Computing ETA depends on numerous parameters. Estimating the arrival time requires combining a large number of static and real-time (dynamic) data sources using state of the art data processing techniques. A further challenge is the decentralised nature of the transport infrastructure which makes prediction cumbersome [87]

In this scenario, the user is a dispatcher that deploys a software that applies advanced routing and stop sequence optimisation algorithm to plan routes and schedules for several vehicles or at fleet scale. The estimated time of arrival of a vehicle can be predicted by the Transport Cloud using static and real-time data. Part of the necessary data is provided by the user (despatcher) about its transport operations (e.g., driver shift time). Through telematics, vehicles provide feedback on their location and progress in their tour (e.g., about which customers have been served). Part of the data is mostly user-independent and collected through a data-sharing platform from a variety of sources.

Computing ETA further enables various subset of services/functionalities such as alerts for delayed stops to the driver; (semi-)automatic update of the tour plan; sharing the arrival time with customers (planning dock availability etc.); rest time planning of the driver; and post-trip reporting and analysis (what causes missed delivery time windows, how narrow can delivery time windows be set, etc.).

[87] D2.9 Use cases definition (V1), 2021, p.16.

Funded by the European Union

These diverse applications with different purposes and data types invoke various regulatory instruments and thus have a multitude of legal implications.

It is observed that currently the market is dominated by a few big players who provide maps and real-time navigation. The data may not be readily available or may be offered through an API at varying prices and conditions.[88] The Transport Cloud will enable the aggregation of data from different platforms while allowing each platform to maintain its business model. This may be seen as a response to the lack of innovative approaches which combine different data sources to obtain more accurate and realistic ETAs.

Taking into account that the part of the data is offered through APIs at varying prices and conditions by a few big players, models and mechanisms which allow aggregating and consolidating data from different platforms are vital for optimisation purposes. Due to the commercial significance of the data for some stakeholders, data pool as a model is not particularly suitable for this type of service. Data pools do not offer sufficient mechanisms to administer different rights and commercial interests but rather focus on making the most data available for all to use without rigid controls and strict tracking of data. The primarily commercial nature of the services renders communal approaches rather unsuitable in this type of service. Diverse commercial interests and possible involvement of personal data do not allow for joint control and decision-making. For similar reasons, data cooperative is also not the proper option—as this organisational form implies a homogeneity of interests among the members of the cooperative.

It should also be considered that the possible involvement of personal data makes it vital that this type of service provide technical solutions with strong anonymisation, tracking and control capabilities. Moreover, as some stakeholders are only willing to offer data on commercial terms, the governance models for Optimisation of Transport flow and ETA must also contain certain monetisation schemes to ensure wide participation. Although the data commons type of models has more management-related features, the joint or collaborative nature of these models makes them prone to similar misfits with data pools. Following from above, *trust-like* models could be more appropriate as they could sufficient legal guarantees (fiduciary) for the control and tracking of data over the Transport Cloud. Given that the fiduciary duties may be calibrated according to the needs of the stakeholders, trust-type forms are particularly suitable for data aggregation involving both private and public actors. It is also possible that certain types of optimisation and ETA services could be established in such a manner to meet the requirements of data altruism organisations.

2.4.3 Use case 2: Emission Reporting

[88] TomTom, for instance, provides traffic and travel time information but it does not share data. Google Maps provides information on the ETA for several transport modes (car, public transport, airplane, bicycle or walking) depending on the data availability and geographical location. D3.3 Market Gap Analysis Report.

The emission reporting use case is concerned with reducing the environmental impact of mobility and transport/t activities. The motivation behind this type of data use is the fact that concrete action for reducing the environmental footprint can only be taken if the impact can be reported clearly and transparently. The use of transport data for emission reporting purposes aims to gain insights into where there is the greatest potential for reducing emissions by comparing different operational (e.g., planning) choices. By exposing the stakeholders' carbon-print it also creates incentives to reduce the environmental impact.

The environmental impact of mobility and transport activities could be handled in two dimensions. *First,* we may speak of the direct and indirect impact of a particular transport asset. The direct impact relates to the operation of the vehicle such as tail pipe emissions. Indirect emission is the environmental impact of the production of the transport assets and equipment (e.g., carbon footprint resulting from the manufacturing of the vehicle). An accurate estimation of direct and indirect emissions is vital for businesses, policymakers and society at large. The *second* dimension of emission reporting is concerned with the share of emissions caused by one item/person/unit. This is no trivial task as it involves the consideration of empty miles or detours and requires a calculation model to assign emissions to a particular activity.[89] Both types of emissions reporting can be integrated into the Transport Cloud services, such as routing, tour planning, tracking, or ETA so that the emission impact both for future and past transport activities could be estimated alongside other activities. Emission reporting, and measuring environmental impact in general, is a complex task influenced by a huge number of parameters, many of which are not readily available.

The Project focuses on providing established emission models with the correct datasets through data-sharing platforms. In the case of direct emission reporting, together with the emission models, the necessary data will consist of tour plans, routes, telematics data, weather, traffic, and vehicle data (e.g., fuel and engine, type, vehicle load etc.).

Currently, mapping and routing companies such as ESRI, HERE, TomTom, Waze, Moovit dominate this domain together with analytics service providers (e.g., Geotab, Inrix, Populus, PTV).[90] Data is generally offered either through an API or a subscription scheme. While the emission reporting of personal cars is well covered by the current services, data for other transport modes is scarcely available. One of the major challenges is different calculation models which makes integration and comparison difficult especially for indirect emissions which involve many parameters at play. The Transport Cloud aims to integrate and create emission-related datasets and provide a catalogue of models for emission modelling. The project output will help close knowledge gaps and for managing related crises through enhanced mitigation, preparedness, response and recovery actions.

The upcoming review of the INSPIRE Directive further intends to increase the availability and reuse of geodata and environmental data. This aims to facilitate the transition to a greener and carbon-neutral economy and reduce the administrative burden for EU public authorities, businesses and citizens. The Project will contribute to the European Green Deal objectives by improving the understanding of governments, businesses and individuals of the societal and environmental impact

[89] D2.9 Use cases definition (V1), 2021.
[90] D3.3 Market Gap Analysis Report, 2021.

of transport and mobility activities. A further aim is to incentivize the use of relevant private sector data to address the climate, biodiversity, pollution and natural resource challenges.

Given that emission reporting requires certain curation and pre-processing of data and the approximation of the predictive models, commons-based governance models may be distinguished as better-suited options. Emissions reporting as an activity could also be situated within a data altruism organisation as they are established to pursue goals that are of "general interest". In this respect, emissions are an exemplary case for the use of data for altruistic purposes. This could be an organisation that collects emissions data from organisations and individuals to further make it publicly available or provide it upon request. A data altruism organisation would be a not-for-profit entity that collects and shares data to improve the quality of life for all. In this regard, the Transport Cloud may contain a data repository where data is shared and accessed for purposes concerning the betterment of society This could also serve as a platform that allows for the easy exchange of data between organisations, researchers, and citizens.

2.4.4 Use case 3: Re-use of transport data for journey planners / digital services

Journey planners and journey planning capabilities are attractive tools that many businesses and digital service providers are willing to integrate into their service offerings. In most cases, these services also include other types of functionalities such as city maps, travel optimisation, vehicle sharing options, retail or tourist attraction applications and catering recommendations.

Currently, it is cumbersome for small businesses and start-ups to handle raw transport datasets as multi-modal journey planning requires combining various types of dynamic and static public transport[91] , geospatial[92], road traffic and ride-sharing data in different formats and standards. Vehicle location data, cartographic data and static infrastructure data are the most commonly available while the payment, ticketing, environment, and dynamic infrastructure data are less easily accessible. This makes integrating data from third parties essential for more efficient journey planning. Although several industries (e.g. health care, tourism and real estate) could be interested in the planning data produced from transport operations, data interoperability legal restrictions the lack of available tools prevent other innovators from accessing planning data.[93]

The Transport Cloud aims to address the need for a common solution for journey planning by offering a service layer (unified API) that simplifies the use for non-domain experts. It is planned that the Transport Cloud will offer specific "on-demand" mobility information, addressed to specific actors

[91] Transportation line, schedules, stop points, stop areas, disruptions, traffic alerts, next arrivals and departures, vehicle information (occupancy, location).
[92] Cartography, addresses, points of interests.
[93] D3.3 Market Gap Analysis Report, 2021.

through a common service layer (e.g., a unified API) based on the Transmodel concepts and data structure.[94]

Regarding suitable data governance models, services limited to providing data analytics functionalities (which do not aggregate data itself) could be administered by simpler models. Yet, planning services that aggregate data require models with more efficient administrative capabilities to manage various types of dynamic and static public transport, geo-spatial, road traffic and ride-sharing data in different formats and standards. Also considering the need for integration data from third parties, efficient journey planning entails a hybrid governance approach combining features from various models. [95] In addition to strong technical support necessary for standardisation and interoperability, the system should also be equipped with certain monetisation tools implementing a data market approach. For journey planning, collaborative models could be combined with other models and organised in both centralised or decentralised ways.[96] The diversity of the stakeholders and data contributors make the membership-based cooperative structure less attractive. Strong technical support and administrative requirements bring to the fore data trusts as they usually contain mechanisms for the administration of access to data and rights clearance.

2.4.5 Use case 4: Analytics and Learning

This use case is about developing general data analysis and learning methods that could contribute to most of the services and activities over the Transport Cloud. This is an important dimension of the Transport Cloud that it also serves as a development platform addressing the needs of municipalities, transport planners or other institutions which seek innovative ways that allow them to offer more efficient, fair and environmentally friendly services.

The objective of this use case is to allow different actors to access and analyse different types of data hosted by the Transport Cloud to extract useful information. The key actors are listed as the researcher, data scientist, and domain expert (data clients) who provide data and carry out analytics to be offered to the decision-maker.[97] The Transport Cloud will contain a metadata catalogue that allows finding, browsing, and exploring datasets that are of interest to the data user. A data API and a service API will provide access to datasets that do and do not require pre-processing respectively.

---

[94] Navitia (https://www.navitia.io/) is an open source trip planner that proposes an open API already based on Transmodel concepts, and therefore could be a good starting point for defining such a standard API that could then be implemented by any journey planner system. Open Trip Planner (OTP, https://www.opentripplanner.org/) is another open source trip planner, that proposes an open API.

[95] D3.3 Market Gap Analysis Report, 2021.

[96] Elinor Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge University Press 1990); Shkabatur (n 55). Michael J Madison, Brett M Frischmann and Katherine J Strandburg, 'Constructing Commons in the Cultural Environment' [2010] Cornell Law Review 657.

[97] D2.9 Use cases definition (V1), 2021.

Funded by the European Union

This use case heavily relies on established (open source) tools that provide many interoperable interfaces to interact with a large pool of data sources. Although there exist analytical tools (software suits) which could be deployed without the need for expert knowledge of ML, these services usually do not offer any data. The holders of the necessary data are generally reluctant to share it relying on IP rights and the control they exercise by owning the infrastructure.

Data altruism organisations provide a suitable legal form to open up transport and mobility data to broader use for purposes aiming for the betterment of human life. Such organisations may work with transportation agencies and other data holders to collect and share data in a way that is useful for researchers, urban planners, developers and others who want to use the data to improve transportation and mobility. The organisation could also develop APIs and other tools to make the data more accessible while also working with other organisations to develop standards for sharing data and create a clearinghouse for data on transportation and mobility. A data altruism organisation to that effect could further be combined with elements from data pools and data commons. Data pools are typically used by organizations to store data that is used internally. Data commons, on the other hand, are often used by communities of users to store data that is used externally.

3. Legal entitlements on data affecting data transactions

3.1 Introduction

Data ecosystems/spaces utilise contracts and transactions to provide access and share data.[98] There is no official definition of a data transaction. Drawing inspiration from contract law, a data transaction (contract) could be defined as any *legal exchange* or *act* which involves either the actual *use of data* and/or *the rights permitting or enabling such use*. A data contract may contain different types of data (e.g., raw data, inferred data, or data protected by Intellectual Property (IP) rights). Similarly, the envisaged use may range from simple access to a permanent or temporary transfer of data (with or without further rights to modify, aggregate, share or otherwise commercialise data). Chapter 2 has provided an account of various governance mechanisms and models that are put in place to effectuate data contracts in several contexts or sectors.

Data contracts may be complex in the sense that they may (expressly or impliedly) incorporate several other transactions (e.g., consent, specific IP license, or other authorisations) which may be subject to different legal regimes and contractual restrictions. Whether that will happen depends on the purpose and the type of data use. It is therefore possible that the use of data as contemplated in a data contract may involve several legal acts and authorisations that may be accompanied by the relevant technical tools.

[98] The present section is an adaptation of the frameworks elaborated in D3.8 (also by KU Leuven) for H2020 Project EUH4D.

The present chapter examines the legal regimes affecting data transactions. Section 3.2 deals with the legal regime for the protection of personal data**,** primarily the GDPR. Section 3.3 focuses on IP rights as a framework conferring exclusive rights on certain types of data which serves as the legal basis for a variety of data exchanges. Section 3.4 analyses the applicability of the TS Directive to data. All these three legal regimes - personal data protection, IP rights and TS protection - are frequently highlighted in the legislative proposals concerning data as legal frameworks that need to be respected. Interestingly, it could be argued that they can both hinder and facilitate data transactions, depending on the nature of the transaction as well as the type of data and the purpose that the data is sought for.

It should be noted that the analysis of the legal regimes is not exhaustive as there might be other legal regimes that could determine the legal limits of data transactions. The analysis rather seeks to provide a view of the legal frameworks that include substantive rights underlying data sharing and reuse.

3.2 Data transactions involving personal data

Rights and obligations relating to the processing of personal data can function both as an enabler and a barrier to data transactions. On the one hand, the general principles (e.g. transparency, data integrity, accountability) and the legal grounds for processing (e.g. consent, see further below in 4.2.1.1) provided in the GDPR[99] restrict data transactions involving personal data in the sense that they create a burden of compliance for the entity mainly responsible for data processing (i.e. the *data controller*) which they might not be able to adhere to. On the other hand, certain individual rights in the GDPR could facilitate data transactions as they allow data subjects to retain some control over their data, thereby providing an incentive to enter into the transaction in the first place. We analyse those rights below, giving special focus on the *right to data portability*[100] . But first, we look into the legal implications of the withdrawal of consent – which is generally relied upon as the legal basis for personal data transactions –where the data subject has already entered into contractual commitments regarding the use of his/her data.[101]

3.2.1 Consent and data transactions

Article 6 of the GDPR provides the legal bases under which personal data could be lawfully processed.[102] But data-driven activities mostly rely upon consent as a legal basis to collect and

[99] GDPR, Art. 5 and 6.
[100] GDPR, Art. 20.
[101] GDPR Art. 7(3) provides that the "*data subject shall have the right to withdraw his or her consent at any time*."
[102] Those are: a) consent, b) necessity for the performance of a contract to which the data subject is party, c) necessity for compliance with a legal obligation to which the controller is subject, d) necessity to protect the vital interests of the data subject or of another natural person, e) necessity for the performance of a task carried

process personal data lawfully.[103] Article 4(11) of the GDPR provides that consent must be "*given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data*".[104]

The notion of 'freely given' suggests that consent should not be forced or pressured from the data subject, while 'informed' suggests that it should be based on facts. In that sense, the data subject's consent is an act of informational self-determination.[105] In cases of apparent imbalance of power between the parties (e.g. contracting with a big tech platform), consent could not be deemed to be given freely.[106] As the requirement of informed consent is to ensure that data subjects will not be deceived or coerced and thereby wronged, it could facilitate data transactions by building trust between businesses and data subjects.[107]

It should be noted that it falls upon the controller to prove that valid consent for a specified purpose was obtained from the data subject. At the same time, consent does not relieve the data controller from the duty of compliance with the general data protection principles such as fairness, necessity, proportionality, and purpose limitation.[108] Unfortunately, in practice, many actors in the data economy treat consent as a '*carte blanche'* to process personal data, even if that processing would be illegal under the GDPR.

Consent, however, has a "weak spot", in that being revocable by the data subject makes data transactions prone to invalidation. Indeed, the data subject has the right to withdraw his or her consent at any time. Withdrawal of consent requires an unambiguous indication of the data subject's will by a statement or clear affirmative action. Article 7(3) GDPR also clarifies that withdrawing consent will not affect the lawfulness of processing based on consent before its withdrawal.

According to the prevalent view (also of the EDPB), even in the case of a contractual setting—when

---

out in the public interest or in the exercise of official authority vested in the controller, f) the controller's or a third party's legitimate interests.

[103] Eleni Kosta, *Consent in European Data Protection Law* (Martinus Nijhoff Publishers 2013).

[104] When processing "sensitive data" (i.e., personal data revealing among other things ethical origin, political opinions, or data concerning the health status of a person or revealing his or her sexual life and orientation), Article 9(2)(a) of the GDPR further requires that the consent must also be explicit.

[105] Antoinette Rouvroy and Yves Poullet, 'The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy' [2009] Reinventing Data Protection: Proceedings of the International Conference (Brussels, 12-13 October 2007) 45; John Kleinig, 'The Nature of Consent' in Franklin Miller and Alan Wertheimer, *The Ethics of Consent: Theory and Practice* (Oxford University Press 2009) 18, 20; Tom L Beauchamp, 'Autonomy and Consent' in Franklin Miller and Alan Wertheimer, *The Ethics of Consent: Theory and Practice* (Oxford University Press 2009); Kosta (n 103).

[106] GDPR, Art. 7 and recital 32.

[107] Laurens Naudts, 'The Right Not to Be Subject to Automated Decision-Making: The Role of Explicit Consent.' (*CITIP Blog*, 2 August 2016) <https://www.law.kuleuven.be/citip/blog/the-right-not-to-be-subject-to-automated-decision-making-the-role-of-explicit-consent/> accessed 12 March 2022; Aurelia Tamò-Larrieux , 'Privacy and Data Protection Regulation in Europe' in Aurelia Tamò-Larrieux, *Designing for Privacy and its Legal Framework*, vol 40 (Springer International Publishing 2018) 89.

[108] Katja de Vries and others, 'Fundamental Rights Protection by Design for Online Social Networks - v2: Update of Deliverable 3' <https://www.usemp.eu/wp-content/uploads/2016/03/usemp_deliverable_d3.6_revised.pdf> accessed 20 April 2022. Also see Article 29 Working Party, 'Opinion 06/2014 on the Notion of Legitimate Interests of the Data Controller under Article 7 of Directive 95/46/EC (WP 217)' 13 <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf> accessed 20 April 2022.

the provision of personal data is a "counter-performance"— the data subject's right to withdraw consent is not affected.[109] Considering the absolute nature of the right to withdraw consent, the possibility that the data subject may exercise his/her right anytime after concluding a data contract creates significant uncertainty which may result in the collapse of a chain of data transactions upon which the businesses rely upon for their operations.

However, it is not possible to categorically conclude that consent is always revocable. Considering that personal data is frequently exchanged as counter-performance in several information society services, such dogma is being contested.[110] There are, to an increasing extent, calls for a more balanced approach by placing certain limitations on the revocability of consent and in general on data subjects' rights (e.g., the right to be forgotten, "RtbF") which may destabilise transactions in the data economy. [111] It is evident therefore that a viable solution that could provide for cohabitation of consent with the contractual dealings on personal data emerges as a pressing issue for the smooth and efficient operation of data markets. The second iteration of this Deliverable (D2.8: "Data Governance recommendations") will further explore the legal solutions to mitigate the uncertainties arising from the inalienable and irrevocable nature of the right to withdraw consent and thus, bring recommendations enhancing the stability of data contracts involving personal data.

3.2.2 Other legal grounds for processing

Aside from consent, other grounds can also be called upon by controllers. Contractual necessity as laid out in the GDPR[112] could, in limited circumstances, be the legal ground for a data transaction. Nevertheless, this is limited to specific contracts and does not constitute a ground for more general processing of personal data. Where certain processing is an indispensable part of a performance or formation of a contract, the data controller may process personal data within this capacity. Under the provision, the performance of a contract may not be made dependent upon the consent to process further personal data which is not needed for the performance of that contract.

Other grounds—legitimate interest of the controller or third parties, protection of "vital interests of the data subject" or the necessity of performance of a task in the public interest—may occasionally be the legal basis for processing personal data without the data subject's consent. Cumulatively, the legal grounds for processing reflect the understanding that lawfulness of data processing requires a

[109] European Data Protection Board, 'Statement 05/2021 on the Data Governance Act in Light of the Legislative Developments' <https://edpb.europa.eu/system/files/2021-05/edpb_statementondga_19052021_en_0.pdf> accessed 20 April 2022. Also see European Data Protection Board and European Data Protection Supervisor, 'EDPB-EDPS Joint Opinion 03/2021 on the Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act)' <https://edpb.europa.eu/system/files/2021-03/edpb-edps_joint_opinion_dga_en.pdf> accessed 20 April 2022.

[110] Swiss Federal Court decision BGE, 136 III 401, May 27, 2010.

[111] Eugenia Politou, Efthimios Alepis and Constantinos Patsakis, 'Forgetting Personal Data and Revoking Consent under the GDPR: Challenges and Proposed Solutions' (2018) 4 Journal of Cybersecurity.

[112] Article 6(1)(b) provides that processing is lawful if "processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract".

balancing of interests among the data subject, data controller and public at large.[113]

3.2.3 Data portability and other data subjects' rights

The GDPR provides an improved set of rights enhancing individual choice, control over data and participative agency.[114] Namely, the right to be forgotten -RtbF- (erasure)[115], the prohibition of automated decisions[116], the right to access and information[117] and the data portability right – which is the main focus of this section - are the most prominent examples of these formulations that are directly relevant to data transactions.[118]

The *data portability* right seeks to facilitate the free movement of personal data. In practical terms, it essentially obliges the data controller to transfer the personal data to the data subject or directly to a third-party of the data subject's choice where such transfer is technically feasible. Data portability provides "a roundabout *business-to-consumer-to-business* (B2C2B) way to achieve B2B data sharing in the case of personal data."[119] Data subjects can have access to their data and avoid lock-in situations where they are tied to one service provider. Conversely, they can 'take' their data and 'move' to another service provider, thereby also stimulating competition between service providers.

Yet, certain limitations exist to the data portability right, preventing full data reuse. The first limitation concerns the personal data and the processing to which it applies. Data portability is only applicable to personal data processing carried out by automated means and based on consent or contractual necessity.[120] Moreover, only the categories of *volunteered* (actively and knowingly provided by the data subject) and *observed* data (left behind by the data subject as a result of the use of a service or device) are covered by the right to data portability. The third category *derived or inferred* data— which is created by the analysis of provided or observed data— is regarded as not covered by the right. The Art.29 Working Party *Guidelines on Data Portability* limit the scope of the right to the

[113] Tamò-Larrieux (n 107) 89–90.
[114] Lazaro C and Le Métayer D, 'The Control over Personal Data: True Remedy or Fairy Tale ?' (Inria - Research Centre Grenoble – Rhône-Alpes ; INRIA 2015) RR-8681 https://hal.inria.fr/hal-01141461, 1 accessed 20 April 2022.
[115] GDPR, Art.17.
[116] GDPR, Art. 22.
[117] GDPR, Art. 15.
[118] Helena Ursic, 'The Failure of Control Rights in the Big Data Era: Does a Holistic Approach Offer a Solution?' in Mor Bakhoum and others (eds), *Personal Data in Competition, Consumer Protection and Intellectual Property Law*, vol 28 (Springer Berlin Heidelberg 2018).
[119] Martens B, de Streel A, Graef I, Tombal T and Duch-Brown N, 'Business to Business Data Sharing: An Economic and Legal Analysis, Digital Economy Working Paper 2020-05, European Commission, Seville, 2020, JRC121336.' (2020) https://ec.europa.eu/jrc/sites/default/files/jrc121336.pdf, 40 accessed 14 March 2022.
; Josef Drexl, 'Designing Competitive Markets for Industrial Data - Between Propertisation and Access' [2017] Journal of Intellectual Property, Information Technology and Electronic Commerce Law 257. Also see Article 29 Working Party, 'Guidelines on the Right to "Data Portability" (Wp242rev.01)' <https://ec.europa.eu/newsroom/article29/items/611233/en> accessed 14 March 2022.
[120] GDPR, Art. 20(1).

Funded by the European Union

personal data provided by the data subject or observed by the controller and thus exclude derived data.[121]

The second limitation concerns technical aspects. Indeed, the GDPR does not stipulate any rules about the process interoperability or technical compatibility between different data controllers' systems. This arguably prevents the provision from being sufficiently operational in the B2B context. The protection of other data subjects and third-party rights also set a global limit to data portability, that is, the arrangements for data portability should not result in unlawful processing of others' data and should not infringe IP rights or TS.[122]

3.3 IP Rights in Data governance

IP rights protect intangible assets such as inventions, business processes and data which are key elements of the digital economy. Different types of rights exist (e.g.., patents, trademarks, designs, copyright and neighbouring rights) but the underlying aim remains the same: to provide an exclusive right that will permit to foster innovation and dissemination of knowledge.

From the perspective of data transactions, the most relevant IP rights are *copyright (author's right)* and the *sui generis* database right. The individual data items (text, audio, video etc.) contained in a database may be eligible for copyright protection under the InfoSoc Directive.[123] Databases will most likely enjoy the *sui generis* protection under the Database Directive[124], which is consequently the major IP regime relevant for data transactions (see further information in section 3.3.2). The copyright protection on databases, as compilations, is rather infrequent since it requires creativity in the selection and the arrangement of the items comprising the database.[125]

As granting exclusive property rights, IP is heavily relied upon to give effect to data transactions. Subject to certain statutory requirements and limitations, IP rights confer exclusive rights in intangible "goods" (informational elements), permitting or denying access to them. IP rights owners can also determine the way and the duration of their usage. As such, IP rights make it possible to exercise control over the informational elements after their release or initial dissemination. Such control constitutes the legal ground for data transactions enabling access, transfer, analysis, or adaptation of data and/or databases.

---

[121] Article 29 Working Party, 'Guidelines on the Right to "Data Portability" (Wp242rev.01)' (n 119) 9–11. However, regarding the question which of the data subject rights (and to what extent) apply to derived data, WP29 Guidelines on Automated Individual Decision-Making and Profiling mentions the possibility of application of other individual rights such as right of access, the right to rectification and erasure to inferred data. For more, see Paul De Hert and others, 'The Right to Data Portability in the GDPR: Towards User-Centric Interoperability of Digital Services' (2018) 34 Computer Law & Security Review 193.

[122] Martens and others (n 119) 42.

[123] Directive 2001/29/EC of the European Parliament and of the Council on the harmonisation of certain aspects of copyright and related rights in the information society [2001] OJ L 167/10 ('InfoSoc Directive').

[124] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases ('Database Directive'), OJ L 77, 27.3.1996, p. 20–28.

[125] Database Directive, Art. 3(1).

3.3.1 Copyright protection of data

Copyright law confers the creators a set of exclusive rights for certain types of use (exploitation) of their works. These so-called *economic*[126] and *moral*[127] rights, preclude third parties from engaging in activities conflicting with the statutorily defined uses of the work. For instance, one may attempt to prevent the use of data in accordance with the contract on economic rights claiming that the data modifications under the contract are prejudicial to her reputation.[128] The economic rights as enshrined in EU legislation include the following rights: to reproduce (copy, rent), communicate and make available to the public and distribute.[129]

Data may be subject to copyright. This concerns mostly the human-created elements (e.g, text, images, videos, sound recordings etc) contained in a data corpus. Under the EU *acquis*, this protection is laid out in the InfoSoc Directive. The copyright on human-created elements may be owned by several different actors in the data governance ecosystem.[130]

Copyright does not protect ideas. Conversely, it protects expressive forms that are original (i.e. 'the author's own intellectual creation') and does not extend to the information contained in a work. Therefore, data relating to factual information either provided by the user or captured through

[126] The *economic rights* may be seen as the variations of a right to use, a right to provide to others, and a right to authorize others to exercise these rights. As exclusive rights, they enable the exploitation of the economic opportunities arising from the monopoly granted by law and therefore, generally accord well with the entitlements stemming from property law. See further in Alexandra George, *Constructing Intellectual Property* (Cambridge University Press 2012) 239.

[127] Moral rights provide creators with control over the treatment and the presentation of their work to others. As formulated under Article 6*bis* of the Berne Convention for the Protection of Literary and Artistic Works, it involves the rights to claim authorship and to object to certain modifications and derogatory actions which would be prejudicial to the author's honour or reputation. As independent rights, they are not attached to the economic rights and thus normally not transferred or cleared in a data transaction. Apparently, moral rights may impair the sustainability of data transactions involving copyright-eligible works.

[128] See Julien Debussche, Jasmien César and Isis De Moortel, 'Leveraging Big Data For Managing Transport Operations (LEMO), Report on Legal Issues, Deliverable D2.2, Horizon 2020 Research and Innovation Programme' (2018) 152 <https://static1.squarespace.com/static/59f9cdc2692ebebde4c43010/t/5bdab3e2cd8366e9378d02b1/1541059569380/D2.2_Report+on+Legal+Issues_LeMO+-+FINAL.pdf> accessed 14 March 2022.

[129] InfoSoc Directive, Articles 2,3,4. As such, the entitlements and empowerments under economic rights, to a certain extent, correspond with the uses of data defined in a given data transaction (e.g., the transfer, copy, modification, analysis and access). A data transaction may involve several aspects relating to somehow exploitation of copyright-protected works and thus, the exercise of economic rights. See further in *Alexandra George* (n 126).

[130] When created by individuals in the course of their engagement with the online services, such material is referred to as User-generated Content (UGC). In many cases, UCG is a derivative of the existing works, merging different materials subject to third-party copyright. For more, see Daniel Gervais, 'The Tangled Web of UGC: Making Copyright Sense of User-Generated Content' (2009) 11 Vanderbilt Journal of Entertainment and Technology Law 841.

sensing or tracking technologies are not copyright-eligible as they lack originality.[131] Any data or information obtained or inferred as a result of data processing (analysis)—be it predictions, personal profiles, or credit scores— do not in itself give rise to a copyrightable work either, as being merely abstract information.[132] Subjective inferences as to the output of data analysis, even in the form of creative opinions, do not amount to a copyrightable form either.

Among the economic rights, the right of reproduction is of major significance to enable data sharing and reuse. Both in international treaties[133] and the InfoSoc Directive, the right of reproduction is understood in a broad sense to include every act of "copying" either in digital or physical form — irrespective of its economic or functional significance.[134] Economic rights such as the right of distribution, broadcast, making available to the public come into play only when the data transaction contemplates any further dissemination of the data eligible to copyright.

In principle, mere access to information or data does not give rise to copyright infringement. However, where a data transaction requires reproduction (e.g., copying, transfer) or adaptation (e.g., transformation, structuring) of data, the exclusive rights of the copyright holder become relevant.[135]

Considering the statutory exceptions and limitations to copyright protection under the InfoSoc Directive, Article 5(1), which allows for acts of temporary reproduction, is subject to narrow interpretation. To benefit from this exception, the temporary copy should be transient or incidental as an integral and essential part of a technological process—aiming solely either enabling transmission in a network between third parties by an intermediary; or a *lawful use* of a work or protected subject matter. The possible applicability of this exception might differ according to the data usage contemplated in a particular data transaction.[136] In particular, the data processing techniques which do not require permanent reproduction such as cloud solutions providing temporary and limited access may benefit from this exception.

[131] It is generally accepted that data such as raw numbers and other purely quantitative information; measurements results (e.g., measurements of temperature, pressure, etc.); financial results, prices of products and similar market data are not eligible to copyright.

[132] What is referred to here as "output data" is also information in machine-usable form. Baskarada S and Koronios A, 'Data, Information, Knowledge, Wisdom (DIKW): A Semiotic Theoretical and Empirical Exploration of the Hierarchy and Its Quality Dimension' (2013) 18 Australasian Journal of Information Systems <http://journal.acs.org.au/index.php/ajis/article/view/748> accessed 14 March 2022.

[133] Berne Convention, 1996 WIPO Copyright Treaty, TRIPS Agreement.

[134] Recital 21 of the InfoSoc Directive. Also see the judgments of CJEU Case C-5/08 (Case C-5/08 *Infopaq International* [2009] ECR I-06569) and; Case C-403/08 (Case C-403/08 *Football Association Premier League and Others* [2009] ECR I-09083) which give the right of reproduction under the InfoSoc Directive an extensive meaning with a view to ensure legal certainty within the internal market.

[135] As data transactions may also involve switching between data formats or selection of certain data from the rest of the corpus, the right of adaptation (Art. 12 of the Berne Convention) may also come into play. Although right of adaptation (as an economic right) is not expressly provided by the InfoSoc Directive and thus not harmonised at the EU level, it is generally regarded to be implicit in the right of reproduction. Christophe Geiger, Giancarlo Frosio and Oleksandr Bulayenko, 'The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market - Legal Aspects' <https://papers.ssrn.com/abstract=3160586> accessed 14 March 2022, see fn 12.

[136] As will be seen below (4.3.2), such exception is not available for sui generis database protection.

The InfoSoc Directive provides for several other optional exceptions or limitations that the Member States may choose to implement to the extent they deem necessary. As a result, the scope of exceptions differs largely within the Union.[137]

3.3.2 Sui Generis Database Right

3.3.2.1 Eligible databases

The Database Directive provides for a s*ui generis* database right as a special type of protection recognized on the entirety of a compilation of data as a systematically produced collection.[138]

The *sui generis* right only protects the database as a collection and does not extend to individual data items contained in the database.[139] This protection, unique to the EU, is without regard to any creativity either as to the content or to the selection or arrangement of the database.[140] The Database Directive does not intend to grant property rights on individual data items but protects the investment in the database as a whole by prohibiting the extraction or the re-utilisation of a substantial part of its contents.[141]

A database is defined as a collection of independent works, data or other material arranged systematically or methodically, and individually accessible by electronic or other means.[142] Irrespective of their copyright eligibility, databases protected by *sui generis* right may consist of any sort of material and in any form whether electronic, paper, online, or hybrid.

---

[137] More on this, see Jean Paul Triaille, Jérôme de Meeûs D'Argenteuil and Amélie de Francquen, 'Study on the Legal Framework of Text and Data Mining (TDM)' (De Wolf & Partners 2014) 31 <https://perma.cc/FXF3-RNWC> accessed 14 March 2022.
[138] Database Directive, Art.7.
[139] Drexl, 'Designing Competitive Markets for Industrial Data - Between Propertisation and Access' (n 119) 269.
[140] For a history and critique of the Directive, see P Bernt Hugenholtz, 'Something Completely Different: Europe's Sui Generis Database Right' in Susy Frankel and Daniel J Gervais (eds), *The internet and the emerging importance of new forms of intellectual property* (Wolters Kluwer 2016).
[141] Francesco Banterle, 'The Interface Between Data Protection and IP Law: The Case of Trade Secrets and the Database Sui Generis Right in Marketing Operations, and the Ownership of Raw Data in Big Data Analysis' in Mor Bakhoum and others (eds), *Personal Data in Competition, Consumer Protection and Intellectual Property Law*, vol 28 (Springer Berlin Heidelberg 2018) 435 <http://link.springer.com/10.1007/978-3-662-57646-5_16> accessed 14 March 2022.
[142] Database Directive, Art. 1.

Funded by the European Union

Extraction, as defined in Article 7(2a) of the Database Directive, could be relevant for data transactions. It refers to permanent or temporary and direct or indirect transfer of contents by any means or in any form. Data transactions aiming for transfer or making available of data may entail the extraction of all or a substantial part of a database. This also applies to indirect or incremental ways which lead to the reconstitution of at least a substantial part of the database. The right of re-utilization in Article 7(2b) may also be relevant. It deals with the dissemination of databases by way of distribution of copies, renting, or other forms of transmission. These two rights entitle the maker of the database to prevent the extraction (akin to reproduction) and re-utilisation (akin to making available) of the whole or a substantial part of the contents of a database. Article 7(5) of the Database Directive prohibits repeated and systematic extractions of a database aiming at reconstituting the whole or a substantial part of the contents of a database.

As the Directive requires a systematic or methodical arrangement, it is subject to debate what type of processing and structuring render raw data eligible for *sui generis* protection. In machine learning ("ML") analysis and big data operations, data may go through intense pre-processing and transformation, and accordingly, qualify for *sui generis* protection as an organized set. However, where the system uses real-time data or unstructured data such as books, pieces of music or video, such corpus may be excluded from protection for not being systematically organised.[143]

3.3.2.2 Substantial investment

Protection under the Database Directive further requires a qualitatively and/or quantitatively substantial investment in either *obtaining, verifying* or *presenting* the contents of the database. It is still not clear how "obtaining" data should be understood. This was elaborated by the CJEU in a series of judgments in 2004 establishing the so-called *spin-off* doctrine.[144] The Court found that the databases such as football fixtures or horse race bulletins did not deserve protection under the Directive as they were the by-products of the main activities of the data controllers, namely organising the horse races and the football league. According to the Court, where the 'creation' of data and the subsequent database is a by-product of the database maker's main activity, such database shall not be protected by the *sui generis* right.[145] The court has limited the application of the doctrine in its later judgment *Football Dataco* by ruling that the facts collected about a football game such as the score, scorer, or penalty decisions were not 'created' data and thus eligible for protection.

Yet, the CJEU has not provided sufficient guidance on how the machine-generated data could be situated within the spectrum between the purely synthetic data and the data observed, therefore we

[143] The question—whether or not such corpus qualifies for sui generis protection— does not in principle affect the possible copyright protection on the individual data items (e.g., audio, video or text).
[144] Case C-46/02 *Fixtures Marketing Ltd v. Oy Veikkaus Ab* [2004] ECR I-10365; Case C-203/02 *The British Horseracing Board Ltd and Others v. William Hill Organization Ltd* [2004] ECR I-10415; Case C-338/02 *Fixtures Marketing Ltd v. Svenska Spel AB* [2004] ECR I-10497; Case C-444/02 *Fixtures Marketing Ltd v. Organismos prognostikon agonon podosfairou AE (OPAP)* [2004] ECR I-10549.
[145] "It is generally understood that, in the case of spin-off databases, companies would have produced these databases anyway without further incentives." European Commission, 'Evaluation of Directive 96/9/EC on the Legal Protection of Databases' (Commission Staff Working Document) SWD(2018) 147 final, 24.

cannot argue in favour of a *per se* inclusion of machine-generated (sensor, IoT) data.[146] Hence it remained questionable whether, under the spin-off doctrine, the operators of ML-based systems can always satisfy the substantial investment requirement in terms of obtaining, presentation or verification of the machine-generated output.[147] This ambiguous position of machine-generated data is now expected to be partially resolved by the European Commission's DA proposal. The data access right granted to the users of IoT devices under Article 4 of the proposed Act provides a limited exemption for machine-generated data from the application of the *sui generis* right.[148]

3.3.2.3 Exceptions and limitations

The Database Directive contains no mandatory exceptions to the *sui generis* right. Nevertheless, to prevent excessive data lock-ups, many limitations, which may be relevant for data transactions, are provided in Articles 8 and 9.

Article 8(1) of the Database Directive, provides that the lawful user of a database, which is made available to the public, is allowed to extract and/or re-utilize the insubstantial parts of its contents. This act therefore may not be prohibited through user agreements or license contracts. [149]

Again, however, it is not evident what "substantial" covers. For instance, a personal profile generated as the output of the ML process may qualify as an independent database. Yet, it may also be a trivial part of a larger database containing all profiles. Based on the way the data is organised and stored, substantiality may depend on the individual circumstances of the case.[150] It was clarified in the BHB case that while assessing substantiality, rather than the intrinsic value of the part used, the substantial investment that relates to the extraction should be given consideration. It could be noted

[146] Some commentators find this approach as severely limiting the application of the Directive, for instance, in the Internet of Things (IoT) environment. Graef argues that spin-off will not be applicable to the "inferred data" accumulating in the hands of the online platforms. See Inge Graef, 'Market Definition and Market Power in Data: The Case of Online Platforms' (2015) 38 World Competition: Law and Economics Review 484.

[147] Matthias Leistner, 'Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform' in Sebastian Lohsse, Reiner Schulze and Dirk Staudenmayer (eds), *Trading Data in the Digital Economy: Legal Concepts and Tools* (Nomos Verlagsgesellschaft mbH & Co KG 2017).).

[148] Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act), COM(2022) 68 final, 1. The provisions and the impact of the Data Act proposal is discussed below in section 4.2.

[149] In the *Ryanair* case the CJEU ruled (Case C-30/14 *Ryanair Ltd v. PR Aviation* [2015] ECLI:EU:C:2015:10 [n 165]) that, since the Database Directive does not apply to databases which do not qualify for sui generis protection, parties were free to determine the contractual conditions of use for databases which are out of the scope of the Directive. Therefore, for databases which do not qualify for the sui generis database protection, Directive's Articles 6(1), 8 and 15 (which preclude contractual limitations) did not apply. The judgment gave rise to a paradoxical situation in that the databases which do not qualify for the sui generis database right received stronger protection through contracts. As presuming an *ab initio* "right" on data, the Court's approach was criticized for contradicting with the rationale of the Database Directive, see Maurizio Borghi and Stavroula Karapapa, 'Contractual Restrictions on Lawful Use of Information: Sole-Source Databases Protected by the Back Door?' (2013) 37 European Intellectual Property Review 505.

[150] Drexl, 'Data Access and Control in the Era of Connected Devices' (The European Consumer Organisation BEUC 2018),) 74.

that the CJEU has generally interpreted the provision broadly in line with the policy objectives of the Directive as protecting the investment of database makers.[151]

The text and data mining (TDM) exception in the Directive on Copyright in the Digital Single Market

The Directive on Copyright in the Digital Single Market (DSM Directive), introduces two mandatory restrictions on copyright and *sui generis* database rights for *text and data mining* (TDM).[152] Since the Directive defines TDM in a way to include a great variety of ML-based analytics, the provisions are likely to be of relevance to data transactions and data governance at large.

The exceptions provided for TDM are without prejudice to the existing exceptions and limitations explained above and Recital 9 further clarifies that the analysis of mere facts or data that are not protected by copyright does not need authorisation. The exception transfers a core principle of copyright into the digital era that factual information remains in the public domain and further encourage the generation of new knowledge which would otherwise not exist due to prohibitive transaction costs. [153]

The unharmonized state of the research and education related exceptions of the EU copyright regime do not fit well with the emerging big data practices. Taking this into account, the Article 3 of the DSM Directive provides an exception relating to research and education activities. With a limited scope, it only covers i) the reproduction of works[154]; ii) temporary or permanent reproduction relating to copyright in the selection and arrangement of databases[155]; iii) extraction and re-utilization of the databases protected by the *sui generis* right[156]; and iv) the press publishers' right in Art. 15(1) of the DSM Directive [157]. Accordingly, research organisations and cultural heritage institutions have an exception in relation to TDM of works or other subject-matter to which they have *lawful access* for scientific research. Article 7 prohibits any contractual provision contrary to the exceptions provided in Article 3.

The other limitation is provided in Article 4 of the DSM Directive. Article 4 provides a general exception allowing TDM on *lawfully accessible* works and other subject matter for any purpose.[158]

---

[151] See cases Case C-545/07 *Apis-Hristovich EOOD v Lakorda AD* [2009] ECR I-01627; Case C-304/07 *Directmedia Publishing GmbH v Albert-Ludwigs-Universität Freiburg* [2008] ECR I-07565.

[152] Directive 2019/790/EC of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Council Directives (EC) 96/9 and 2001/29 [2019] OJ L 130/92 ('DSM Directive'). Articles 3 and 4. see Daniel Gervais, 'Exploring the Interfaces Between Big Data and Intellectual Property Law' (2019) 10 JIPITEC 3, paras 38–40

[153] Benjamin Raue, 'Free Flow of Data? The Friction Between the Commission's European Data Economy Initiative and the Proposed Directive on Copyright in the Digital Single Market' (2018) 49 International Review of Intellectual Property and Competition Law 379, 381.

[154] InfoSoc Directive, Art. 2.

[155] Database Directive, Art. 5(a).

[156] Database Directive, Art. 7(1).

[157] The provision obliges news aggregators who link to publishers' content or use snippets, e.g., Google news, to obtain a license.

[158] Unlike Article 3, Article 4 additionally covers the permanent or temporary reproduction of computer programs where the normal use of the program necessitates such reproduction. The reason for this

However, the provision only applies to certain exclusive rights in a fragmented way and therefore it is far from being a general non-infringement exception.[159] But more importantly, Article 4 is not applicable if the rightsholders bring reservations in an "appropriate manner" to restrict TDM. In cases where the content is made publicly available online, only machine-readable (code-based) measures are considered appropriate (Recital 18). For other content, contractual arrangements and unilateral declarations are also acceptable means to circumvent the TDM exception.

Concerning data transactions, the exceptions provided in the Directive might be of relevance as having both facilitative and restrictive effects depending on the context. The TDM exception may enable data transactions as it may empower the parties for certain specific uses of data owned by a third party. Yet, the broad allowance for the contractual circumvention in Article 4 is the major shortcoming of the provision which is likely to render it partially inefficient.

3.4 Trade secret protection and data

3.4.1 Introduction

Data or datasets may constitute or contain valuable or sensitive information that requires protection against unlawful acquisition or disclosure. Aside from the IP toolbox, TS protection may be applicable. TS allows businesses to control information that is not eligible for other types of IP protection.[160]

Contrary to the different IP rights analysed below, no specific categories exist defining the subject matter eligible for TS protection. A trade secret can be any information of business value, which is kept secret and which somehow provides an economic advantage to its holder. Combined with technological protection measures and contractual arrangements, this broad scope of protectable information makes trade secrets the preferred "appropriability mechanism" for many businesses.[161]

As such, TS protection effectively addresses the need for confidentiality for it legally protects physical secrecy even where copyright protection is unavailable or ineffective. Yet, TS protection is not a substitute for IP rights but may rather be used in a complementary fashion.[162] Both TS protection and the sui generis rights may subsist in the same dataset.[163]

---

discrepancy between Art. 3 and 4, leaving computer programs out of the scope of the scientific research exception, is unclear.

[159] Gervais (n 152) paras 44–45.

[160] Robert G Bone, 'The (Still) Shaky Foundations of Trade Secret Law' (2014) 92 Texas Law Review 1803.

[161] European Commission, 'Study on Trade Secrets and Confidential Business Information in the Internal Market (Final Study, Contract Number: MARKT/2011/128/D, 2013)'.

[162] Brenda Simon and Ted Sichelman, 'Data-Generating Patents' (2017) 111 Northwestern University Law Review 377, 389.

[163] See European Commission, 'Evaluation of Directive 96/9/EC on the Legal Protection of Databases' (Commission Staff Working Document) SWD(2018) 147 final (n 145) 43.

Similar to IP rights explained above, in the context of data transactions, trade secrets also come into play under two scenarios, namely TS owned by one or more of the transacting parties and TS claims by third parties. As such, TS protection too may act both as a barrier and a facilitator for data reuse and sharing.

3.4.2 EU Trade Secrets Directive

At the EU level, TS protection is harmonised by the Trade Secrets Directive.[164] The Directive provides a minimum standard of protection and refrains from obliging the Member States to recognize property-based rights over information. It rather lays out a framework allowing the Member States to maintain their preferred type of protection, as long as undisclosed know-how and business information are safeguarded against misappropriation.[165] The formulation used in the Directive follows the wording of the TRIPs Agreement – the international framework for IP protection. Accordingly, unlawful acquisition of a trade secret means *"unauthorised access to, appropriation of, or copying of any documents, objects, materials, substances or electronic files, lawfully under the control of the trade secret holder, containing the trade secret or from which the trade secret can be deduced"* (Article 4(2a)). Acquisition of trade secrets through other types of conduct that are considered contrary to honest commercial practices is also prohibited (Article 4(2b)).

3.4.3 Trade secret protection of data

Individual data items contained in a dataset could satisfy the secrecy requirement under Article 2(1a) of the TS Directive, in the sense of not being generally known or readily accessible. Commonly available factual information such as one's age, gender etc. will difficulty meet the secrecy threshold. On the contrary, information such as the exact location of a pothole on the city roads (as is known by many of the citizens) poses a more difficult question—not lending itself to a straightforward answer.[166]

TS law grants legal protection to *de facto* secrecy. The requirement of reasonable steps to keep information secret could be achieved both by technical and organizational measures such as digital encryption, designating restricted areas in the company premises or introducing individual access restrictions. Businesses also heavily make use of contractual clauses mandating confidentiality or

[164] Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ L157/1 ('Trade Secrets Directive') .
[165] Gintarė Surblytė, 'Enhancing TRIPS: Trade Secrets and Reverse Engineering' in Hanns Ullrich and others (eds), TRIPS plus 20: From Trade Rules to Market Principles, vol 25 (Springer Berlin Heidelberg 2016) 726; Drexl, 'Data Access and Control in the Era of Connected Devices' (n 150) 91.
[166] Josef Drexl and others, 'Data Ownership and Access to Data - Position Statement of the Max Planck Institute for Innovation and Competition of 16 August 2016 on the Current European Debate' (Max Planck Institute for Innovation and Competition 2016) para 21 <www.ip.mpg.de/en/link/positionpaper-data- 2016-08-16.html> accessed 14 March 2022.

prohibiting reverse-engineering.[167] The adequacy of the measures is relative and will be determined in consideration of the TS holder's economic size, sectoral conditions, prior experience, organisational policies and so on. In this regard, the reference made to reasonableness and specific circumstances in Article 2(1c) may be regarded to pinpoint a test of proportionality.[168]

In the case of analysis of real-time data, there could be special difficulties in the enforcement of the secrecy measures arising out of the intricate supply chains and numerous participants which make contractual arrangements impractical or too costly. In such cases, system owners and operators also resort to technical solutions such as APIs to keep their data secret while allowing for limited use.[169]

Leaving this aside, data generated, for instance, by the heat sensors in a machine would qualify as TS for containing valuable information about the manufacturing process.[170] Subject to limitations of the privacy and data protection regime, information derived from personal data could also enjoy TS protection.

Other than individual data items, databases may also enjoy TS protection. The source of the data—whether it is obtained from individuals, measured by sensors, generated in a machine-to-machine process or captured through tracking technologies— does not have a bearing on the evaluation of the secrecy requirement.[171] According to Article 2(1) of the TS Directive, in terms of the secrecy requirement, a database must be treated as a unit in its entirety. The information must be secret "in the sense that it is not, as a body or in the precise configuration and assembly of its components, generally known". For instance, the aggregated customer database as a whole may well qualify as a TS, even if the information about individual customers can be found in open sources. The secrecy with respect to a database also includes how individual pieces of information relate to each other.[172] Accordingly, a personal profile comprising of many selected, possibly trivial factual data may possess the necessary quality of confidentiality.[173] This also applies to datasets enabling the deduction of information protected by TS.[174]

Datasets could be treated confidential both due to the information they contain and also for their function in a certain process. For instance, in the case of a properly labelled training dataset (used for developing ML models), the TS protection on such data does not necessarily aim to maintain the confidentiality of certain information but rather to deprive the competitors of a useful tool. Hence, the

[167] Mariateresa Maggiolino, 'EU Trade Secrets Law and Algorithmic Transparency (Bocconi Legal Studies Research Paper No. 3363178)' (2019) 9.
[168] Robert G Bone, 'Trade Secrecy, Innovation and the Requirement of Reasonable Secrecy Precautions' in Rochelle C Dreyfuss and Katherine J Strandburg (eds), The Law and Theory of Trade Secrecy (Edward Elgar Publishing 2011) <http://www.elgaronline.com/view/9781847208996.xml> accessed 14 March 2022.
[169] Dag Wiese Schartum, 'Making Privacy by Design Operative' (2016) 24 International Journal of Law and Information Technology 151.
[170] Herbert Zech, 'A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data' (2016) 11 Journal of Intellectual Property Law & Practice 460.
[171] On trade secret protection of the data gathered via the Internet of Things, see Cristiana Sappa, 'What Does Trade Secrecy Have to Do with the Interconnection-Based Paradigm of the Internet of Things?' (2018) n°8-40 European intellectual property review 518.
[172] Drexl, 'Data Access and Control in the Era of Connected Devices' (n 150) 93.
[173] Drexl and others (n 166) para 25.
[174] Josef Drexl, 'Legal Challenges of the Changing Role of Personal and Non-Personal Data in the Data Economy' in A De Franceschi and R Schulze (eds), *Digital Revolution: Data Protection, Smart Products, Blockchain Technology and Bitcoins Challenges for Law in Practice* (Beck 2019).

training data should rather be considered as a kind of *resource*, distinct from the concrete semantic information it embodies.[175]

Regarding the commercial value requirement in Article 2(1b), the Directive does not set a threshold but rather deems commercial value implicit, based on the investment made for obtaining or generating data or for the effort to keep it secret. Even publicly available data, seemingly without any commercial value, may provide a competitive advantage when compiled into a database. As Recital 14 of the TS Directive states that value could be actual or potential, unstructured data could also qualify as TS.[176] In sum, the existence of a market for the data may be seen as the prima facie evidence of its commercial significance.

3.4.4 Restrictions on trade secret protection

The TS Directive does not grant a property right in the information but rather establishes a liability regime. The protection is not unconditional but subject to TS holders' strict preservation of the *de facto* secrecy. A TS will cease to exist even when the information is made public via illegitimate means without the consent of the right holder.[177] The same applies if rival businesses or in general third parties discover the undisclosed information through independent efforts, by way of inspection or reverse-engineering (which is not prohibited). Under Article 3(1b) of the TS Directive, the observation, study, disassembly or testing of a product or object that has been made available to the public or that is *lawfully in the possession of the acquirer* is permissible.

However, the discovery of a TS by way of reverse engineering is only possible if there is no 'legally valid duty' to the contrary (e.g., a contract clause prohibiting reverse engineering). This allowance of contractual restrictions under the TS Directive severely diminishes the benefits from reverse-engineering as a limitation to TS protection. [178]

[175] Drexl, 'Designing Competitive Markets for Industrial Data - Between Propertisation and Access' (n 119) 281, para 127.
[176] N Sousa e Silva, 'What Exactly Is a Trade Secret under the Proposed Directive?' (2014) 9 Journal of Intellectual Property Law & Practice 923, 924; Drexl, 'Designing Competitive Markets for Industrial Data - Between Propertisation and Access' (n 119) 269, para 54.
[177] Maggiolino (n 167) 11. This is irrespective of any monetary compensation for damages that may arise from the unlawful appropriation or disclosure of the secreted information.
[178] David D Friedman, William M Landes and Richard A Posner, 'Some Economics of Trade Secret Law' (1991) 5 Journal of Economic Perspectives 61, 62–63.

Funded by the European Union

4. The emerging EU data governance regime

4.1 The European Data Strategy and the common European mobility data space

The uncertainties and imbalances stemming from the current fragmented regulatory landscape (where IP rights, trade secrets and personal data protection rules apply concurrently) have been a major concern underlying numerous EU policy instruments and legislative initiatives.[179] To address the current shortcomings regarding the sharing and reuse of data, in its 2020 Data Strategy, the Commission identifies several critical issues that need to be overcome to foster the availability of data, eliminate imbalances in market power, ensure data interoperability and empower individuals to exercise their rights. Based on the fact that a high degree of market power enables Big-Tech companies to impose unilateral conditions for access to and use of data, The strategy document draws attention to the imbalances relating to access to co-generated IoT and industrial data which give rise to disadvantages in terms of developing new services and products.[180]

Under the emerging EU data governance regime (essential pillars of which have been laid out by the 2020 Strategy document) the organisational, transactional and technical dimensions of data governance are structured around the concept of "data space". The creation of EU-wide common, interoperable data spaces in strategic sectors aims at overcoming legal and technical barriers to data access and sharing—bringing together relevant data infrastructures and governance frameworks. Key features of the common European data spaces are secure and privacy-preserving infrastructure to pool, access, share, process and use data; clear and practical structure for access to and use of data in a fair, transparent, proportionate and/non-discriminatory manner; respect for European rules and values; openness; and data control.

Initially, ten data spaces are planned with the possibility of additional ones, ultimately creating a European data space as a genuine single market for data.[181] Among them, the common *European mobility data space* aims to accelerate the digital transformation of the transport sector. It builds upon the existing EU-wide and domestic legal, technical and organisational initiatives. In the area of transport and mobility, there exist various current and prospective regulatory instruments which organise data access and sharing in B2B, B2G, G2B and G2G contexts. Many of these instruments implement certain architecture, platform or governance tools for the harmonisation of data-sharing and access conditions.

[179] European Commission, 'Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A Digital Single Market Strategy for Europe COM(2015) 192 Final'; For an account of the development of this agenda, see Sebastian Lohsse, Reiner Schulze and Dirk Staudenmayer (eds), *Trading Data in the Digital Economy: Legal Concepts and Tools* (1st edition, Nomos ; Hart Publishing 2017).
[180] European Commission, 'A European Strategy for Data ' (n 39) 8.
[181] European Commission document on Common European Data Spaces (n 34), 1.

Under the *Sustainable and Smart Mobility Strategy*[182], future legislative initiatives are planned for the access and sharing of mobility data aiming to contribute to the development of the mobility data space. In the air traffic sector, the amended version of the proposal for a Regulation on the Single European Sky[183] includes provisions on data availability and market access of data service providers in the field of air traffic management. Type approval legislation[184] in the automotive sector (currently under review), contains rules for the access of third-party service providers to data relevant to repair and maintenance information. To modernise the EU's transport system and support the transition to cleaner, greener and smarter mobility, the Commission intends to update the 2010 *Intelligent Transport System Directive* proposing that certain crucial road, travel and traffic data (e.g., speed limits and traffic circulation plan) are made available in digital format.

The Commission further plans revisions of the Delegated Regulation (EU) 2017/1926 *on multimodal travel information services* to include mandatory accessibility to dynamic datasets; the Directive 2005/44/EC on *harmonised river information services* (RIS) to enhance the integration of inland waterway transport into multimodal logistics, contributing to the interoperability of information services and data sharing; and the technical specifications for interoperability for telematics applications for rail passengers. The Commission also plans to propose rules on a trusted environment for the corridor data exchange framework to support collaborative logistics, based on recommendations from the *Digital Transport and Logistics Forum* (DTLF).[185]

The European Commission has also announced its support for the deployment of the common *European mobility data space* by providing funding to various projects under the Digital Europe Programme (DEP)[186] for the creation of technical infrastructures and governance mechanisms to facilitate easy, cross-border access to key data resources, This will be accompanied with a preparatory action for a comprehensive mapping of the existing initiatives to propose concrete actions aiming for harmonisation and interoperability. There are also initiatives for the construction and establishment of various data ecosystems, platforms and marketplaces led by the Member States or private actors.[187]

4.2 The Data Act proposal

The 2020 Strategy document sets up an action plan which envisages two main legislative proposals as the foundation of the upcoming EU data governance regime. The first proposal, the DGA

[182] COM(2020) 789 final.
[183] COM (2020) 579 final.
[184] Regulation (EU) 2018/858 of the European Parliament and of the Council REGULATION (EU) 2018/858 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 30 May 2018 on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles, amending Regulations (EC) No 715/2007 and (EC) No 595/2009 and repealing Directive 2007/46/EC.
[185] European Commission document on Common European Data Spaces (n 34) , 17-18.
[186] Digital Europe Programme, Work Programme for 2021-2022, https://ec.europa.eu/growth/sectors/tourism/funding-guide/digital-europe-programme_en.
[187] European Commission document on Common European Data Spaces (n 34) , 18-19.

(published on the 25th of November 2020) complements and further solidifies the emerging regime relating to public sector data.[188] The DGA proposal aims to make public sector data available for local businesses, researchers and communities for the development of innovative data-driven services on a larger scale. It has a specific focus on the public sector data which is subject to legal restrictions and thus left out of the scope of the Open Data Directive. The DGA proposal covers public sector data that is legally protected on the grounds of (a) c*ommercial confidentiality including the trade secrets*; (b) statistical *confidentiality*; (c) *intellectual property* rights of third parties; (d) *protection of personal data*. The public sector bodies enabling the use of such protected data are required to be technically equipped to ensure that data privacy and confidentiality are fully preserved. The proposal does not interfere with the substantive rights on data as it refrains from prescribing a right of *access* or *reuse* but lays out certain harmonized rules and conditions guiding Member States in establishing mechanisms for the reuse of publicly held data. Considering the objectives of the MobiDatalab project and the services offered on the Transport Cloud, it is of note that the framework contemplated under the DGA requires that public sector bodies should avoid contracts that grant exclusive rights to data unless such exclusivity is necessary for the provision of a service or a product in the general interest.

The second most important legislative initiative under the 2020 Data Strategy is the Data Act (DA) which aims to encourage and enable greater and fair B2B and B2G data use in all sectors.[189] The proposal which was released on 23.02.2022[190] is regarded to be an essential building block of the European data spaces.[191] It is guided by the understanding that B2B contractual agreements do not fully guarantee adequate access to data for SMEs or start-ups—entailing a contractual framework providing clarity as to the rights and remedies regarding accessing, processing, sharing and storing of data to limit the misuse of such data. The proposal acknowledges the importance of a harmonised data governance regime in achieving competitiveness, innovation and sustainable growth in all sectors and making the Union's transition to a green digital economy a success.

The proposal introduces several interventions to the existing regulatory and contractual framework regarding data reuse and sharing both in the B2B and B2G contexts. First, in terms of B2B data use, the DA proposal provides data access rules for the users (and to third parties at the request of the users) of IoT products or related services (Chapter II, jointly with Chapter X as an enabler). Second, Chapter III is aimed at constituting a *lex generalis* for data sharing obligations to be laid down in the future. There is a direct – although implicit – connection with data spaces, for which more specific (and especially, concerning mobility, sector-specific) data sharing obligations could be laid down in the future. Third, the proposal harmonises the rules to prevent the exploitation of contractual imbalances between businesses by mainstreaming the principle of fairness in B2B commercial data transactions, although only to the benefit of SMEs, which significantly reduces the scope of application (Chapter IV), facilitate switching between data processing services (Chapter VI) and enhance interoperability (Chapter VIII). In the B2G context, the DA proposal in Chapter V provides for the making available of data held by private entities to public sector bodies in exceptional

[188] The proposal defines 'public sector body' as: the State, regional or local authorities, bodies governed by public law or associations formed by one or more such authorities or one or more such bodies governed by public law.
[189] European Parliament resolution of 25 March 2021 on a European strategy for data (2020/2217(INI)) OJ L 77, 27.3.1996, p. 20–28
[190] Proposal for a Regulation on harmonized rules on fair access to and use of data (Data Act), COM(2022)68 final ('Data Act proposal').
[191] European Commission document on Common European Data Spaces (n 34), p.2.

situations. As seen, these interventions come in varying scales and scope and interacts with numerous legislative instruments.

The Data Act regulates cloud and edge services in several respects. Chapter VI addresses lock-in vendor issues left unsolved by the Free-Flow of Non-Personal Data Regulation[192]. Similar to the provisions of the DGA concerning public sector bodies, data altruism organizations and data intermediaries[193], Chapter VII aims to preserve the European Union 'digital sovereignty' by laying down safeguards applicable when a foreign law would require international transfer of non-personal data from cloud and edge services providers. Art. 29 lays down the essential requirements for interoperability for cloud and edge services, subject to further regulation by the European Commission.

Below we will provide a short commentary of the rules introduced by the DA proposal regarding the sharing and reuse of data.

4.2.1 Data access right

The idea of a *data access right* is motivated by the proliferation in the products and related services equipped with sensors, cameras, microphones, gyroscopes, radar, and similar functionalities to collect data, and by the subsequent need to clarify who is entitled to use data generated by such objects and on what basis. The aim is to create favourable conditions in particular for small businesses and start-ups by offering innovative and sustainable solutions in all sectors of the economy. The proposal aligns with the objectives of EU consumer protection law as it provides more transparency and the ability to access data generated from the use of an IoT product.

Chapter II of the DA proposal provides rules relating to access to data both in B2B and B2C contexts. *Data access right* applies only to physical, movable products that obtain, generate or collect, through their physical components, data concerning their performance, use or environment and that can communicate that data via a publicly available electronic communications service (often referred to collectively as the IoT).[194] This also covers the data generated by related services the absence of which would prevent the product from performing its functions.[195] Related services need not be directly provided by the seller, renter or lessor but could be offered by a third party.[196] Data access right specifically covers virtual assistants which serve as an interface to reach content or activate IoT objects— acting as a gateway to record significant amounts of relevant data on how users interact with the products. Considering that the design requirement may turn out to be burdensome for micro- and small enterprises, Chapter II does not apply to data generated by the use of products manufactured or related services provided by micro or small enterprises (Article 7(1)).

[192] Regulation (EU) 2018/1807 of 14 November 2018 on a framework for the free flow of non-personal data in the European Union, OJ L 303/59 ('Free-Flow of Non-Personal Data Regulation'). For an analysis of the Regulation, please see D2.1.
[193] On this topic, see Baloup and others (n 32).
[194] Such products may include vehicles, home equipment and consumer goods, medical and health devices or agricultural and industrial machinery, see DA proposal, recital 14.
[195] DA proposal, Article 2 (2) and (3).
[196] DA proposal, recital 17.

Funded by the European Union

Chapter II provides two types of obligations addressed to manufacturers and designers on the one hand and data holders on the other. Device manufacturers are in a position to determine (through their control of the technical design of the product or related services) what data are generated and how they can be accessed—including the cases where they have no legal entitlement on the data. Article 3 puts manufacturers and developers under an obligation to design their products in a way that the data is easily accessible in a transparent manner and to provide the users with the relevant information.[197,198] It is further noted that the product should not be designed to permit the manufacturer to monitor the activities of the user or a third party, e.g., when or how often the data is accessed or processed. Such monitoring could, inter alia, give the manufacturer an unfair competitive advantage. The provision may be seen as a specific application of the concept of *legal protection by design* and therefore, it may be seen as complementary to *data protection by design* (DPbD) provided in the GDPR.

Article 4 provides the obligation on the *data holders*[199] to make available to the user the data generated by the use of an IoT product or a related service—without undue delay, free of charge and, where applicable, continuously and in real-time. The user or a third party is precluded from exploiting the data obtained under Article 4 to develop a product that competes with the product from which the data originate. It is also made clear that the data holder may only use any non-personal data generated by the use of a product or related service based on a contractual agreement with the user.[200] Data holders are precluded from using this data to derive insights about the economic situation, assets and production methods of or the use by the user in a way undermining the commercial position of the user (Article 4(6)).

Under Article 5, data holders are further obliged to make available the data generated by the use of a product or related service to a third party upon the request of the user. This obligation shall be fulfilled without undue delay, free of charge to the user and of the same quality as is available to the data holder and, where applicable, continuously and in real-time. The third party is only permitted to process the data for the purposes agreed with the user and may only share it with another party in case the service requested by the user necessitates such further sharing. The third-party shall not deploy coercive means or abuse evident gaps in the technical infrastructure of the data holder designed to protect the data. (Article 5(4)). Reciprocally, the data holder is also prevented from using any non-personal data generated by the use of the product or related service to derive insights about the economic situation, assets and production methods or in a way that undermines the commercial position of the third party (Article 5(5)).

---

[197] This Regulation (Data Act) is without prejudice to existing and future Union law setting physical design and data requirements for products to be placed on the European Union market.

[198] "Products may be designed to make certain data directly available from an on-device data storage or from a remote server to which the data are communicated. Access to the on-device data storage may be enabled via cable-based or wireless local area networks connected to a publicly available electronic communications service or a mobile network." DA proposal, recital 21.

[199] 'Data holder' means a legal or natural person who has the right or obligation, in accordance with this Regulation, applicable Union law or national legislation implementing Union law, or in the case of non-personal data and through control of the technical design of the product and related services, the ability, to make available certain data. DA proposal, Article 2(6).

[200] The proposed Act does not prevent the data holder from proposing contractual conditions relating to the purchase, rental or lease of the product, whose effect is to allow the data holder itself also to access and use the data.

Articles 4 and 5 specifically provide that all necessary measures should be taken to preserve the confidentiality of trade secrets. Where the requested data involves any personal data generated by the use of a product or related service, the data holder shall not make such data available without a legal basis under Article 6(1) of the GDPR. This applies to both the access right of the user in Article 4 and the third-party access in Article 5.[201]

Article 6 lays out the rules that the third parties obtaining data are required to comply with. As mentioned above, third parties receiving data under Article 5 are required to process the data only for the purposes and under the conditions agreed with the user. In the case of personal data, the third party shall delete the data when they are no longer necessary for the agreed purpose. Third parties shall particularly refrain from using the data for the profiling of natural persons within the meaning of Article 4(4) and 22 of the GDPR; make the data further available to another third party (unless such provision of data is necessary to provide the service requested by the user), and using the data to develop a product that competes with the product from which the accessed data originate. Third parties cannot impose contractual terms on the user preventing the user from making the data available to other parties.[202]

As the most important intervention of the proposed DA, Chapter X, in a single provision (Article 35), provides that the *sui generis* right provided for in Article 7 of the Database Directive (96/9/EC) does not apply to databases containing data obtained from or generated by the use of a product or a related service. This specifically addresses the uncertainties and confusions about the applicability of the sui generis rights in the IoT context as explained in Section 3.2.2 above. Accordingly, concerning the exercise of the right of access in Article 4 and the third-party access in Article 5, the database to which access is sought will not be subject to the *sui generis right* provided by the Database Directive.

4.2.2 General rules applicable to obligations to make data available

Chapter III of the DA proposal sets out general rules relating to the obligations to make data available which either stem from horizontal regulations or vertical interventions in the areas such as banking, vehicles and electricity. According to Article 8, a data holder who is obliged to make data available to a data recipient under Article 5, other Union law or national legislation implementing Union law shall perform this obligation under fair, reasonable and non-discriminatory terms and in a transparent manner under the provisions of the chapters III and IV of the proposed DA.

Chapter III applies only to legally mandated data access or sharing requirements.[203] The onus of proof lies with the data holder where a data recipient contends that the conditions under which data has been made available are discriminatory. Similar to the data access right, Article 8 also refers to trade secrets stating that, in the absence of rules to the contrary under Union law or national

[201] DA proposal, Article 4(5).
[202] DA proposal, Article 6(2).
[203] Data use which is purely based on freedom of contract remains unaffected by these rules.There is also reference to personal data providing that the obligations to make data available under the GDPR were not affected by the rules provided in the proposed Data Act.

legislation implementing Union law, any obligation to make data available shall not oblige the disclosure of trade secrets**.**

Article 9 lays out the conditions under which the data holder may demand reasonable compensation unless further rules exist on this matter (e.g. specified in sectoral legislations). Where the data recipient is a micro, small or medium enterprise[204], the compensation cannot exceed the costs directly related to making the data available. Article 10 establishes a dispute settlement mechanism (through entities certified by the Member States) to settle disputes about the determination of fair, reasonable, transparent and non-discriminatory terms.

Considering the regulatory agenda of the EU commission and also the initiatives regarding the establishment of European Data Spaces, it is understood that more sectoral mandatory data access and data provision requirements will be introduced by the EU legislature and other competent authorities at various levels as explained in Section 4.1 above. An important point to note is that Article 12(3) sets a time limit to the application of the provisions of Chapter III, excluding the obligations to make data available under Union law or national legislation implementing Union law that existed before the entry into force of the proposed DA.[205]

4.2.3 Unfair contractual terms in data sharing

Considering the disproportionate bargaining power between the parties to a data contract, the resulting unfair contract terms particularly harm micro, small and medium-sized businesses as they cannot negotiate the conditions for access to data. The terms in data contracts often put an unfair burden on a start-up or a small company and therefore make access to data commercially less viable or attractive. Such contractual terms restrict the ability of the weaker stakeholders to develop or run innovative data-driven business models. Preventing such effects promotes innovation and ensures a fair allocation of value creation in the data economy.

Chapter IV of the proposed DA provides that unfair contractual terms concerning the access and use of data or the liability and remedies which are unilaterally imposed on micro, small or medium-sized enterprises shall not be binding (Article 13).[206] The provision particularly concerns the situations where one party supplies a contractual term to a weaker party (a micro, small or medium-sized enterprise) who is unable to influence the terms of the contract through negotiation. Article 13 will not apply to the parts of the contract which are not related to making data available, in particular, the contractual terms defining the main subject matter of the contract or determining the price to be paid. By way of reference provided in Article 8(1), the unfairness test provided in Article 13 will also apply to an obligation to make data available as defined in Chapter III.

---

[204] See Article 2 of the Annex to Recommendation 2003/361/EC.
[205] See also Article 40 of the DA proposal.
[206] A contractual term that is simply provided by one party and accepted by the micro, small or medium-sized enterprise or a term that is negotiated and subsequently agreed in an amended way between contracting parties should not be considered as unilaterally imposed.

Under Article 13, paragraphs 3 and 4 respectively lay out the conditions where a contractual term will be held or presumed unfair. Accordingly, contract terms that aim for or result in excluding or limiting the liability of the imposing party intentional acts or gross negligence; excluding the remedies available to the weaker party in case of non-performance of contractual obligations; giving the imposing party the exclusive right to determine conformity of the supplied data or to interpret any term of the contract shall be regarded as unfair and thus will not be binding (Article 13(3)).

The presumption of unfairness applies to the contract terms which inappropriately limit the remedies in case of non-performance of contractual obligations or the liability in case of breach of those obligations; allow the imposing party to access and use data of the other contracting party in a manner that is significantly detrimental to the legitimate interests of the other contracting party; prevent or limit the weaker party from using or obtaining a copy of the data contributed or generated by that party during the period of the contract, or enable the imposing party to terminate the contract with unreasonably short notice (Article 13(4)).

It is made clear that the principle of freedom of contract remains intact as an essential element in B2B relationships, meaning that not all contractual terms happen to be more favourable to one party than to the other but only those that are excessive and abusive clauses shall be struck down. Pursuant to Article 13(5), the onus of proof is on the imposing party to show that the allegedly unfair term has not been unilaterally imposed. As a general principle of contract law, Article 13(6) provides that where an unfair contractual term is severable from the remaining terms of the contract, those remaining terms will be upheld.

4.2.4 Exceptional use by public sector bodies

Chapter V of the DA proposal establishes a framework for the use of the data by public sector bodies and Union institutions where there is an exceptional need. Article 15 defines three cases where an exceptional need shall be deemed to exist: i) responding to a public emergency; ii) preventing a public emergency or assisting the recovery from a public emergency, and iii) fulfilling a specific task in the public interest that has been explicitly provided by law. It is held that, in such cases, the public interest relating to the use of the data outweighs the interests of the data holders.

Public emergencies include various natural or human-induced contingencies such as health or environment-related disasters or cybersecurity attacks. The decision as to the presence of a public emergency shall be determined according to the respective procedures in the Member States or of relevant international organisations.[207] Other than public emergencies, the exceptional need could also be justified where the lack of timely access or the unavailability of data makes the fulfilment of public duties impossible.[208]

Chapter V does not apply to data needs about prevention, investigation, detection or prosecution of criminal and administrative offences, the execution of criminal and administrative penalties or the collection of data for taxation or customs purposes. These activities remain to be subject to their specific regulations. The small and micro enterprises are also exempted from the obligation to make

[207] DA proposal, recital 57.
[208] Ibid, recital 58.

data available as provided in the Chapter.[209] As a further limitation, the Open Data Directive [210] shall not apply to the data made available to public bodies under Chapter V. Hence, the data entrusted to public bodies for exceptional use cannot be considered as open data available for reuse by third parties.[211] Having said that, Recital 68 states that public bodies could share the data with other entities or persons to carry out data analysis (e.g. machine learning) operations that are beyond the capacities of the requesting public body.

The Chapter does not provide an exception to the *sui generis* database right similar to the access right in Article 4. Yet, Recital 63 states that where *sui generis* database right[212] applies to the requested datasets, data holders should exercise their rights in a way that does not prevent the public sector bodies or Union institutions from obtaining the data, or from sharing it in accordance with this Regulation.

4.3 Further provisions and the evaluation of the Data Act

As the second major legislative initiative announced in the 2020 Data Strategy, the DA proposal primarily deals with the B2B data (personal and non-personal) exchanges with a particular focus on the control and monetisation of data cogenerated through IoT products and related services. It aims the creation of a cross-sectoral governance framework for data access and sharing and to this end, it introduces rules that regulate relations between actors in the data economy, incentivising horizontal data sharing.[213] Such legislative intervention has been considered necessary to ensure legal certainty and transparency for economic operators and micro-, small- and medium-sized enterprises.

Other than the provisions explained in the above sections of this Chapter, Chapter VI obliges data processing service providers to take necessary measures to ensure their customers can switch to similar data processing services. This covers removing the commercial, technical, contractual and organisational obstacles that prevent customers from terminating the contract, entering into a new contract for the same type of service; or porting their data, and other *digital assets*[214] to another provider. Article 24 contains detailed rules as to the contractual terms concerning switching between providers of data processing services. Article 26 provides further technical requirements to ensure that, where available, services are compatible with open standards or interfaces. Considering the limited efficacy of the self-regulatory frameworks and shortcomings and limited scope of consumer protection, Chapter III intends to expand the right of data portability in Article 20 of the GDPR to non-personal data.

---

[209] Ibid, Article 14(2).
[210] Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (OJ L 172, 26.6.2019, p. 56).
[211] DA proposal, Article 17(3).
[212] Database Directive (n 124).
[213]DA proposal (n 182) , p.1.
[214] "Digital assets refer to elements in digital format for which the customer has the right of use, including data, applications, virtual machines and other manifestations of virtualisation technologies, such as containers. Functional equivalence means the maintenance of a minimum level of functionality of a service after switching, and should be deemed technically feasible whenever both the originating and the destination data processing services cover (in part or in whole) the same service type. Meta-data, generated by the customer's use of a service, should also be portable pursuant to this Regulation's provisions on switching." DA proposal, recital 72.l.

Chapter VII provides that *international access and transfer* requiring that *data processing services*[215] shall take all reasonable technical, legal and organisational measures, including contractual arrangements, to prevent the unlawful international transfer or governmental access to non-personal data held in the Union. Article 27 also lays out the conditions such as a valid international agreement for the transfer of data to a third country. It is stated that the provisions of the proposed Act comply with the Union's international commitments in the World Trade Organization ("WTO") and other in bilateral trade agreements.[216]

Chapter VIII lays out interoperability requirements addressed to *operators of data spaces* and data processing service providers together with requirements relating to smart contracts. It is of note that the addressees in this Chapter also include data space operators as a novel type of actor.[217] Under Article 28, data space operators are obliged to describe the datasets contained in a specific data space in various dimensions. These requirements include content, use restrictions, licences, collection methodology, quality, structures, formats, vocabularies, classification schemes, taxonomies, and code lists relating to data together with the technical means of access (e.g. application programming interfaces-APIs). Data space operators shall also provide information as to their terms of use and quality of service to enable automatic access and transmission of data between parties. The Commission is empowered to request one or more of the European standardisation organisations to draft harmonised standards that satisfy the essential requirements laid out in Chapter VIII.[218] The Commission shall also adopt common specifications, where harmonised standards are non-existent or insufficient to ensure conformity.[219] In addition, it remains at the discretion of the Commission to further adopt guidelines prescribing interoperability specifications for the functioning of common European data spaces. These may include architectural models and technical standards implementing legal rules and arrangements between parties such as technical translation of consent or permission.[220] Article 29 enables open interoperability specifications and European standards for the interoperability of data processing services to promote a seamless multi-vendor cloud environment.

Article 30 on *smart contracts* also contains requirements for data access and sharing. Smart contracts in the context of an agreement to make data available shall i) present a very high degree of robustness to avoid functional errors and to withstand manipulation by third parties; ii) contain mechanisms to terminate the continued execution of transactions; iii) enable archiving of the transactional data together with the code to ensure audibility; and iv) establish mechanisms to control access at the governance and smart contract layers.

---

[215] "'Data processing service' means a digital service other than an online content service as defined in Article 2(5) of Regulation (EU) 2017/1128, provided to a customer, which enables on-demand administration and broad remote access to a scalable and elastic pool of shareable computing resources of a centralised, distributed or highly distributed nature." DA proposal, Article 2(12). .

[216] DA proposal (n 182) , p.16.

[217] The DA proposal does not provide a definition of the term *"operators of data spaces"*.

[218] Regulation (EU) No 1025/2012 of the European Parliament and of the Council of 25 October 2012 on European standardisation, amending Council Directives 89/686/EEC and 93/15/EEC and Directives 94/9/EC, 94/25/EC, 95/16/EC, 97/23/EC, 98/34/EC, 2004/22/EC, 2007/23/EC, 2009/23/EC and 2009/105/EC of the European Parliament and of the Council and repealing Council Decision 87/95/EEC and Decision No 1673/2006/EC of the European Parliament and of the Council, OJ L 316, 14.11.2012, p. 12–332.

[219] DA proposal, Article 28 (4) and (5).

[220] Ibid, Article 28(6).

Funded by the European Union

Finally, Chapter IX provides an implementation and enforcement framework together with a complaint mechanism. It is stated that the Commission shall recommend voluntary model contractual terms on access to and use of data. The Chapter also foresees penalties that will apply in case of infringement of the provisions of the Regulation (DA).

Overall, the DA proposal may also be seen as a consolidation of the existing rules, remedies and legal affordances scattered in various legal instruments contained in the EU *acquis*. The data access right (though with limited scope), general rules for mandatory data access, and the rules on unfair contract terms may be inferred or derived from the legal frameworks (e.g., relating to consumer protection, unfair contract terms, and general principles of contract law) currently in force in the Union or Member State laws. Yet, the proposed Regulation is an important step in terms of pointing the direction that the EU will follow in regulating and clarifying the legal uncertainties regarding B2B, B2C and B2G data use.

Important to note, neither the provision of the proposal nor the recitals or other explanatory parts of the text provides any methodological clarity or guidance as to the systematics of the Act, in particular the cumulative effect of the provisions in Chapter II, III and IV on data contracts/transactions, and they will interact and complement each other. All three issues handled in the B2B context namely, access by the users to the IoT device data, general rules for the obligations to provide data and the rules relating to unfair contractual terms address data use at different scales. While access right offers a rather solid remedy to the part of the problem, the scope of the general rules in Chapter III and their practical implementation remains vague. The chapters of the prosed Act are addressed to different types of actors, that is, *IoT device manufacturers and related service providers*, *data holders*, *data processing service providers*, *data space operators* and *vendors of an application using smart contracts*. Considering that the groups of stakeholders such *as data space operator* or *smart contract vendor* do not have clear legal definitions, identifying the exact group of stakeholders which are subject to different obligations and the possible overlaps may pose difficulty.

Both the proposal and the preparatory work contain a strong emphasis on the need to balance the policy objectives of IP protection of databases in the context of the data economy. The exemption of machine-generated data from the scope of the sui generis right is a solid step in terms of clarifying uncertainties concerning access to IoT data. Yet, the copyright rules for creative works and the Trade Secret protection remain applicable. The proposal repeatedly refers to trade secrets in various provisions stating that appropriate measures shall be taken to preserve the confidentiality of the trade secrets.[221] However, these references provide almost no guidance about how the provisions of the proposed Act relating to access or sharing of data will be applied to cases where the data question contains or constitutes a trade secret.[222]

Both the recitals and the provisions of the proposed Act address the question of interaction with the GDPR on many occasions. It is generally confirmed that where the data in question wholly or partly qualify as personal data the applicable rules of the GDPR should be complied with. Recital 7 explicitly states that the provisions of the proposed Act could not be applied or interpreted in such a way that would diminish or limit the right to protection of personal data. It is also made clear that the DA does not give rise to a legal basis under the GDPR for the data holder to provide access to personal data when requested by a user that is not the data subject. Hence, it could be concluded that the provisions of the GDPR (as protecting a fundamental right) take precedence over the provisions of

[221] DA proposal, Articles 4(3), 5(9), 8(6), 17(2), 19(2).
[222] See section 3.4.3

the DA. Yet, this does not solve the problem at hand because the proposed DA, like some other legislative instruments, heavily rely on the distinction between personal and non-personal data. The issue is that such distinction increasingly becomes blurred as datasets are often mixed in nature containing both types of data. Moreover, data aggregation from diverse sources together with the deanonymisation techniques enable the extraction of personal information from seemingly non-personal or anonymised data.

4.4 The impact on the MobiDataLab Project

The proposed Act is expected to significantly impact the cloud market and current contractual framework and thus the activities and services contemplated within the MobiDataLab project. Given that in the transport, logistics and mobility sectors innovation and efficiency heavily relies on the sharing and exchange of large amounts of personal and non-personal data between multiple actors, the proposed Act supports and aligns with the objectives of the MobiDataLab project.

The data access right, as being the most concrete intervention together with the exemption of sui generis right provided in the proposal, will be directly applicable to the mobility devices and products (such as rental cars, bikes, navigation devices and other equipment which falls under the definition provided in Article 2 (3) and (4). The access right does not apply to mobile phone applications such as journey planners since personal computers, servers, tablets, smartphones, cameras, webcams, sound recording systems and text scanners are excluded as they require human input to produce various forms of content.[223] Hence the devices connected to the Transport Cloud and the related services will need to be designed in such a way that both the data recorded intentionally by the user (a.k.a voluntary data) and the data generated as a by-product even in stand-by mode or switched-off (a.k.a. observed data) will be accessible by the users of the product or will be made available to a third-party upon request. It is made clear that data access right does not extend to data or information derived or inferred by the product manufacturer or related service provider or by other lawful holders of data.[224]

General rules applicable to obligations to make data available in Chapter III of the DA is of high significance for the mobility and transport sector. As has been explained, the provisions under this Chapter aim to ensure that the conditions for mandatory data access satisfy certain criteria. The Chapter provides a general framework that should be complied with when fulfilling obligations for making data available such as the access right introduced in Article 4 of the DA proposal. However, the extent of the chapter goes beyond the access right and it also covers mandatory data access or data use contemplated under the sectoral rules and regulations such as the Delegated Regulations that specify data accessibility for road and multimodal passenger transport within the framework established by the Intelligent Transport Systems (ITS) Directive.[225] It should be noted that by virtue of Article 12(3) which sets a time limit to the application of the provisions of Chapter III the general rules will apply to future Delegated Regulations under the ITS Directive together with the upcoming legislative and regulatory instruments regarding sectoral data access and provision requirements.[226]

[223] Data generated through these excluded devices could be accessible pursuant to the provision of the GDPR provided that such data qualify as personal data.
[224] DA proposal, recital 14 and 17.
[225] OJ L 207, 06.08.2010, p. 1-13.
[226] See section 4.1 above.

Hence, it is currently unclear whether and how the existing data access and data provision rules enacted pursuant to Intelligent Transport Systems Directive will be handled.

The fairness test in Chapter IV relating to data contracts will oblige those who offer data on the Transport Cloud to observe certain rules in their contractual dealings. Considering that the "data holder" as a legal term is defined in very broad terms including every natural and legal person who legally or physically controls data, those who offer data under contractual terms in the Transport Cloud will need to adjust their contracts in accordance with the fairness test.

In sum, it is clear that the DA together with the DGA make up the main building blocks of the emerging *EU data access regime* with multi-layered rights, obligations and technical and organisational requirements in a multitude of dimensions. Therefore, as concrete effects and implications of this novel regime become more solidified in the coming months, the second iteration of this deliverable will focus on the necessary adjustments and alignments entailed by the proposed DA.[227] Lastly, it should be mentioned that some early critiques argue that DA should have focussed more on incentivising data sharing rather than imposing strict obligations on parties at this very early stage of developing Europe's data economy.[228] Comments from the data holders' side draw attention that the rules seem to be aiming to invert established market dynamics by deepening the regulatory control of a wide variety of data-related transactions— imposing rules that are very cumbersome to monitor and enforce.

[227] D2.8: Data Governance recommendations (v2)
[228] Luca Bertuzzi, Industry readies to fight the Commission's Data Act proposal https://www.euractiv.com/section/digital/news/industry-readies-to-fight-the-commissions-data-act-proposal/

5. Conclusion

The analysis in this document covering the legal, administrative and organisational implications of data governance models/mechanisms and the legal frameworks applicable to data transactions yield results that are of significance to the Project and the Transport Cloud.

The *conceptual framework of data governance* introduced in Chapter II attempts to establish a theoretical basis for a more methodological approach and a comprehensive analysis of the emerging EU data governance regime. As such the conceptual framework in Section 2.2 provides the preliminary input for *a holistic interpretation* of the regulatory corpus (EU *acquis*); and an *efficient implementation* of the upcoming legislation on data access and sharing (e.g., DGA, DA, DSA, DMA).

The analysis of data governance models and mechanisms in Chapter II clearly illustrates the lack of consensus as to the exact nature, characteristics and legal status of these organisational structures. Since various forms, structures and tools are discussed and experimented with, Chapter II by no means provides an exhaustive list of data governance. As seen, in real-life scenarios, illustrated by the use cases relating to specific applications on the Transport Cloud, it is not possible to find clear-cut examples precisely corresponding to the explained models but rather hybrid formulations containing elements from various models and mechanisms. Hence, market structures and business practices in the mobility or transport sector are not mature enough or sufficiently established to be easily associated with a model or mechanism with precise features and uniform application. The analysed organisational models and mechanisms do not coherently fit into a categorisation or taxonomy. What could be concluded about the governance models and mechanisms is that they all present similar legal challenges that can partly be linked to uncertainties relating to privacy, permissible types of data use and technical implementation which act as a deterrent for researchers, investors and initiators.[229] Considering the complexity and the dynamism of the European legal landscape relating to data access and sharing, further refinement and clarity regarding these models emerge as a pressing need that will be handled in the second iteration of this Deliverable (D2.8: "Data Governance recommendations").

Chapter III, in section 3.2, reveals that achieving a balance between individuals' data privacy and fostering data access and sharing still requires significant scholarly and practical efforts. A novel approach to GDPR needs to be developed—based on *civil law* and the doctrine of *protection of personality*— to reach a more stable legal regime of data transactions. Hence, D2.8 "Data Governance recommendations" will further explore the legal solutions to mitigate the uncertainties arising from the inalienable and irrevocable nature of the right to withdraw consent and thus, offer recommendations to enhance the stability of data contracts involving personal data. Section 3.3 on IP rights illustrates that the transactions involving data that are subject to copyright and sui generis database right pose significant difficulties in terms of clearance of third-party rights and the transfer of economic rights which are based on rigid statutory categories. As explained, with the exemption provided in Article 35, the differing views and discussions about the application of sui generis right to machine-generated data is partially resolved. Yet, other cases of possible IP application still pose complex problems which could not be easily resolved without legislative intervention. Analysis of the

[229] For instance, data trusts, data commons and other models which require assignment of rights by the members/partners suffer from the legal complications underlying data contracts.

Trade secret (TS) protection in section 3.4 yields similar results. The legal protection of sensitive business information through the TS Directive could foster data transactions by unlocking data that would be kept in the dark in the absence of TS protection. Having said that, extensive reliance on trade secrets could also severely impede the efficient operation of data markets as TS protection covers any type of information. In sum, solutions to practical problems relating to IP rights and trade secrets require the deployment of various regulatory tools and formulations of a plethora of legal instruments, rules, doctrines from various branches of law.

The DA proposal analysed in Chapter IV aims to address the shortcomings resulting from the incomplete and fragmented legal landscape explained in Chapter III. It is seen that the Proposal introduces obligations and defines rights that aim to make data available to a wider range of stakeholders. The Proposal, to a certain extent, temper the existing market order and structures favouring large data incumbents at the expanse of smaller (European) actors. The binding rules obliging making data available to the users of IoT devices together with a fairness test and the general rules have already attracted criticism.[230] The broad territorial reach, alterations to current data sharing and access practices mostly established by the dominant actors, and the extensive technical specifications are expected to generate strong opposition from interest groups and obviously, mobility and transport sectors are no exception. [231]

Finally, it needs to be mentioned that the repeated delays of the DA proposal and the intensity of the ongoing legislative process envisaged under the 2020 European Data Strategy has to some extent prevented an extensive elaboration of the emerging EU data governance regime at this stage. The implementation of the legal solutions and the specifications contemplated by the new legislation will require several micro-arrangements and the synchronisation of many “moving parts”. Hence, it is planned that the effect of DA and DGA on the Project will be elaborated and handled more systematically in the second iteration of this Deliverable.

[230] The views submitted by a diverse range of stakeholders in the public cpnsultation process revealed that a mandatory access regime for machine-generated data or a reallocation of rights on co-generated data was not the preferred solution by the big players. See feedback from B*roadcom, IBM, EuroCommerce* ‘Data Act & Amended Rules on the Legal Protection of Databases - Feedback and Statistics: Inception Impact Assessment’ (*Have your say*) <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13045-Data-Act-&-amended-rules-on-the-legal-protection-of-databases/feedback_en?p_id=24828813> accessed7 April 2022.

[231] Clément Perarnaud and Rosanna Fanni , The EU Data Act: Towards a new European data revolution? Luca Bertuzzi, European Commission pitches data sharing obligations in Data Act proposal, EURACTIV.com 23.02.2022.

MobiDataLab consortium

The consortium of MobiDataLab consists of 10 partners with multidisciplinary and complementary competencies. This includes leading universities, networks and industry sector specialists.

@MobiDataLab
#MobiDataLab

https://www.linkedin.com/company/mobidatalab

For further information please visit **www.mobidatalab.eu**

MobiDataLab is co-funded by the EU under the H2020 Research and Innovation Programme (grant agreement No 101006879).