Cambridge Law Journal, page 1 of 34 © The Author(s), 2026. Published by Cambridge University Press on behalf of The Faculty of Law, University of Cambridge.
doi:10.1017/S0008197326101585

UNLOCKING DATA: A MATRIX OF DATA-SHARING REGULATION IN THE AGE OF AI

ORIT FISCHMAN-AFORI* AND SHARON BAR-ZIV**

ABSTRACT. *This article addresses the challenges in formulating data-sharing regulations by proposing a systematic regulatory matrix based on data characteristics. While data-driven innovation drives economic growth, existing legal frameworks in the EU are incoherent and often tend towards data propertisation, even if indirectly, which may lead to data underutilisation. The matrix is built on varied data characteristics, and it aims to foster access to data and data sharing as a fundamental general principle underpinning the data-driven innovation market. This framework offers a balanced regulatory scheme ranging from open access to proprietary models, aiming to maximise innovation and public good in the emerging field of data law.*

KEYWORDS: *data*, *data sharing*, *data governance*, *data regulation*, *artificial intelligence*, *data taxonomy*, *open access*.

I. INTRODUCTION

Data are increasingly central to the modern economy. The "data products and services" market, comprising various technologies related to data collection, storage and analysis, has prospered in the past two decades.[1] In recent years, this trend has accelerated with the maturation of artificial intelligence ("AI") technologies which rely heavily on data to generate new insights and develop innovative applications.[2] Such data-driven products and services are already in use across a broad range of fields, including, for example, personalised recommendation systems in e-commerce (e.g. after booking a flight, additional services at the destination, such as hotels, are promoted) and various automated decision-making tools used in

*Faculty of Law, College of Management, 7 Rabin Blv. Rishon Le'Zion 7502501 Israel. Authors are listed in alphabetical order and contributed equally to this article.
**School of Law, Sapir Academic College, Mobile Post Hof Ashkelon 7915600 Israel. Address for Correspondence: sharonbz@mail.sapir.ac.il.

[1] European Commission, "European Data Market Study 2022–2023", 11, available at https://digital-strategy.ec.europa.eu/en/library/results-european-data-market-study-2021-2023 (last accessed 27 March 2026).

[2] OECD, "Digital Economy Outlook 2024 (Volume 1): Embracing the Technology Frontier", 38–40, available at https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/05/oecd-digital-economy-outlook-2024-volume-1_d30a04c9/a1689dc5-en.pdf (last accessed 27 March 2026).

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

essential sectors such as finance (e.g. decisions on loan or insurance terms), healthcare (e.g. decisions concerning medical procedures and treatment) and transportation (e.g. decisions made by autonomous vehicles).[3] These developments have contributed to what many describe as the "data economy", in which access to and control over data has become a central driver of commercial value and economic growth.[4]

Recognising that data-driven innovation constitutes a key pillar of contemporary economic growth, the EU adopted the Digital Single Market Strategy in 2015 and the European Strategy for Data in 2020, with the aim of promoting data-driven innovation across sectors.[5] Building on these strategic frameworks, the EU undertook an intensive process of designing different regulations to maximise the potential inherent in data. All these regulations directly or indirectly address various aspects of control over or enable access to data. Yet, the legal frameworks addressing data sharing or access to data are, to some extent, incoherent, often involving a confused discussion regarding data with varying characteristics in diverse contexts. This situation has led to a very complex and convoluted legal framework pertaining to data, with many uncertainties.[6] In response to these shortcomings, the EU proposed a Digital Omnibus in 2025 aimed at simplifying and consolidating several aspects of the existing legal frameworks.[7]

Moreover, existing European regulatory mechanisms are grounded in legal frameworks that frequently give rise to direct or indirect proprietary regimes over data. For example, intellectual property laws may occasionally contribute to the establishment of exclusivity over data, even if only indirectly.[8] Control over data, and thus exclusivity, may also be established through possession of the data-storage device; through technological protection measures such as encryption and access controls

[3] European Commission, "Data Market Study", 68–69.

[4] European Commission, "Shaping Europe's Digital Future", available at https://eufordigital.eu/wp-content/uploads/2020/04/communication-shaping-europes-digital-future-feb2020_en_4.pdf (last accessed 27 March 2026); OECD, "Digital Economy Outlook", 47.

[5] European Commission, "A Digital Single Market Strategy for Europe", available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52015DC0192 (last accessed 27 March 2026); European Commission, "Shaping Europe's Digital Future".

[6] H. Schweitzer and A. Metzger, "Data Access under the Draft Data Act, Competition Law and the DMA: Opening the Data Treasures for Competition and Innovation?" (2023) 72 GRUR International 337, 388.

[7] European Commission, "Proposal for a Regulation [...] Amending Regulations (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725, (EU) 2023/2854 and Directives 2002/58/EC, (EU) 2022/2555 and (EU) 2022/2557 as regards the simplification of the digital legislative framework and repealing Regulations (EU) 2018/1807, (EU) 2019/1150, (EU) 2022/868 and Directive (EU) 2019/1024", available at https://eur-lex.europa.eu/resource.html?uri=cellar:ebf17714-c56e-11f0-8da2-01aa75ed71a1.0001.02/DOC_1&format=PDF (last accessed 27 April 2026) ("Digital Omnibus Proposal").

[8] P.B. Hugenholtz, "Against 'Data Property'" in H. Ullrich, P. Drahos and G. Ghidini (eds.), *Kritika: Essays on Intellectual Property*, vol. 3 (Cheltenham 2018), 48.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

creating "digital locks";[9] or through contractual arrangements that restrict access to data.[10] Additionally, privacy interests are often invoked to limit data sharing, arguably contributing to an exclusionary regime.[11] Collectively, these legal frameworks reflect a clear tendency towards the "propertisation" of data.

A comprehensive analysis of these frameworks reveals a patchwork of rules resulting in inconsistencies regarding whether, and to what extent, data should be subject to a proprietary regime. The fragmented legal framework governing data access has created ambiguity and has arguably restricted data sharing,[12] and is said to contribute to the underutilisation of data for innovation.[13] This reality raises important questions about how data sharing should be regulated as an essential driver of innovation – questions that have become central to the data economy landscape.[14]

This article addresses the challenge of data-sharing regulation by offering a systematic analysis of the concept of data and proposing a regulatory matrix grounded in a guiding principle that may enhance legal coherence. We argue that data governance, and particularly data sharing, requires a foundational theoretical approach that can serve as a determinant principle for recognising exclusivity over data. The European strategy for data regulation explicitly refers to the goals of innovation, efficiency and social welfare, reflecting a utilitarian orientation in policy design.[15] This discourse therefore also engages with traditional theoretical justifications for intellectual property rights, which have been extensively discussed for decades.[16] Within this broad discourse, scholars have increasingly questioned the applicability of

[9] European Parliament and Council Directive 29/2001 (OJ 2001 L 167 p.10), art. 6 ("InfoSoc Directive").

[10] V. Moscon, "Data Access Rules, Copyright and Protection of Technological Protection Measures in the EU: A Wave of Propertisation of Information", available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4515815 (last accessed 27 March 2026); W. Kerber, "Governance of IoT Data: Why the EU Data Act Will Not Fulfill Its Objectives" (2023) 72 GRUR International 120.

[11] N. Purtova, *Property Rights in Personal Data: A European Perspective* (Alphen van den Rijn 2011); J. Crémer, Y.-A. de Montjoye and H. Schweitzer, "Competition Policy for the Digital Era", 77–78, available at https://op.europa.eu/s/Aeuh (last accessed 27 March 2026).

[12] J. Asswad and J. Marx Gómez, "Data Ownership: A Survey" (2021) 12 Information 465.

[13] H. Schweitzer et al., "Data Access and Sharing in Germany and in The EU: Towards a Coherent Legal Framework for the Emerging Data Economy", 70–72, 119, 128, available at https://pure.mpg.de/rest/items/item_3457829_2/component/file_3457831/content (last accessed 27 March 2026); M. Draghi, "The Future of European Competitiveness – Part A: A Competitive Strategy for Europe", 30, available at https://commission.europa.eu/document/download/97e481fd-2dc3-412d-be4c-f152a8232961_en (last accessed 27 March 2026).

[14] OECD, "Enhancing Access to and Sharing of Data in the Age of Artificial Intelligence", 11, available at https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/02/enhancing-access-to-and-sharing-of-data-in-the-age-of-artificial-intelligence_d70ccab9/23a70dca-en.pdf (last accessed 27 March 2026).

[15] European Commission, "Digital Single Market Strategy", 3; European Commission, "Shaping Europe's Digital Future", 2–3.

[16] P. Drahos, *A Philosophy of Intellectual Property* (London 1996); T. Dreier, "Balancing Proprietary and Public Domain Interests: Inside or Outside of Proprietary Rights?" in R. Dreyfuss, D. L. Zimmerman and H. First (eds.), *Expanding the Boundaries of Intellectual Property – Innovation Policy for the Knowledge Society* (Oxford 2001), 295; H. Ullrich, "Intellectual Property: Exclusive Rights for a Purpose – The Case of Technology Protection by Patents and Copyright" in K. Klafkowska Wasniowska (ed.), *Problemy Polskiego I Europejskiego Prawa Prywatnego* (2012), 425; J. Pila, *The Subject Matter of Intellectual Property* (Oxford 2017), 30.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

the incentive-based utilitarian approach to data.[17] Building on this critique, we propose a conceptual framework for understanding data that can serve as a foundation for data-sharing regulation. To serve the public interest in fostering data-driven innovation markets more effectively,[18] we advocate adopting a general default rule that favours broad access to data or the depropertisation of data, while recognising the establishment of exclusive rights only as a justified exception.

The data fuelling the thriving data economy originates in activities across all sectors, private and public, either as a primary goal or as a by-product of their operations. This multifaceted data undoubtedly presents a major challenge for regulators, yet it calls for a structured data governance framework that promotes innovation for the benefit of society. The proposed matrix is built on varying data characteristics and, to construct it, we introduce a taxonomy of data based on three guiding questions:

(1) What are the characteristics of data?
(2) Who collected and stored the data?
(3) Why were the data collected and stored?

We argue that data-sharing regulation should take account of this taxonomy, in which the questions of why the data were collected – whether incidentally or purposefully, for commercial or non-commercial use – and who collected the data should serve as key considerations. Exclusive control over data generated through the operation of digital products and services should be acknowledged only when it is necessary as an incentive for collecting and storing the data, typically when the data have been collected purposefully – that is, intentionally – for commercial purposes and as part of the business model of the collecting entity. In practice, data-collection activities are varied and complex. The proposed matrix is designed to address these contingencies and offers a spectrum of regulatory approaches that can be determined according to the level of merited exclusivity. At one end lies a fully open-access model with unrestricted data sharing; at the other, a proprietary model that grants full exclusivity. Between these extremes lie complex scenarios, which are the most challenging to design. In these cases, under the proposed matrix, the rule should be calibrated to either deviate from the proprietary model by imposing data-sharing obligations or to soften the open-access model by allowing some exclusivity, based on the required incentives, as assessed through the matrix's guiding taxonomy questions.

[17] H. Zech, "Information as Property" (2015) 6 Journal of Intellectual Property, Information Technology and E-Commerce Law 192; W. Kerber, "Specifying and Assigning 'Bundles of Rights' on Data: An Economic Perspective", available at https://www.econstor.eu/bitstream/10419/234876/1/dp2021-22.pdf (last accessed 27 March 2026).
[18] European Commission, "Shaping Europe's Digital Future", 6.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

The significance of this matrix lies in its potential to stimulate the development of a comprehensive and systematic framework for data-sharing regulation. The proposed analysis may serve as a tool for the critical assessment of existing regulations and could provide a foundation for coherent future developments in the emerging field of "data law" (*lex data*), in a manner that maximises the public good by promoting data innovation.[19]

This article proceeds as follows. Part II describes the centrality of data in the age of AI. Part III reviews the frameworks governing data, including traditional frameworks – such as intellectual property rights and possession – as well as newer EU data regulations. This review highlights the ongoing transition from exclusive to access-oriented approaches to data. Part IV discusses the need to shift from an exclusive to an open-access regime, which lies at the heart of the policy matrix, and Part V presents the taxonomy underlying its development. Part VI explains the proposed data-sharing regulation matrix. Part VII discusses the advantages and limitations of the proposal, and Part VIII concludes.

II. Data in the Age of Artificial Intelligence

The information age is characterised by an unprecedented acceleration in data generation and collection. This proliferation stems from a multitude of digital interactions between individuals, electronic devices (e.g. computers, mobile and medical devices) and other entities.[20] Data typically move through a life cycle of generation, collection and storage. Generation occurs, for example, through individuals' activities (such as user online interactions), sensor operations (such as telecommunications signals) or automated processes (such as autonomous vehicle operations); collection brings data under an entity's control; and storage enables their retention over time. However, the distinction between data generation and data collection is not always clear-cut as in many digital systems data are generated and captured simultaneously. Nevertheless, while generation marks the creation of data, control over data is exercised primarily at the stages of collection and storage.[21]

These immense volumes of collected and stored data, commonly referred to as "big data", serve as a significant resource for information processing and knowledge creation, facilitating the development of innovative products, processes and markets.[22] The convergence of extensive

[19] OECD, "Age of Artificial Intelligence".
[20] R. Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructure and Their Consequences* (London 2014).
[21] J.M. Wing, "The Data Life Cycle" (2019) 1 Harvard Data Science Review.
[22] J. Krämer, D. Schnurr and S. Broughton Micova, "The Role of Data for Digital Markets Contestability", 15, available at https://cerre.eu/wp-content/uploads/2020/08/cerre-the_role_of_data_for_digital_markets_contestability_case_studies_and_data_access_remedies-september2020.pdf (last accessed 27 March 2026).

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

databases, advanced computing capabilities and sophisticated algorithms has enabled more comprehensive and nuanced data analysis, contributing to the rapid advancement of data-driven products and services.[23] A substantial leap in data markets has been associated with the emergence of mature AI technologies, which have significantly enhanced the capacity for large-scale data analysis and the generation of novel insights.[24] AI technologies are developed on the basis of large-scale databases and their continued development necessitates the increasing use of extensive and up-to-date data resources.[25] Therefore, in the current age of AI, access to data and its use have become key policy questions.

III. Data Propertisation: A Patchwork of Traditional and New Legal Frameworks

Data are regulated by a wide range of legal frameworks. The evolution of these frameworks reflects the challenges posed by rapid technological development and the various ways in which the law responds to the complex realities associated with data. Within this evolutionary process, traditional legal frameworks have been adapted to address data, alongside newer regulatory initiatives specifically targeting data governance. While none of these frameworks establishes full-fledged ownership in data as such, they acknowledge forms of control over data that generate legally protected exclusivity. Considered together, they reveal a cumulative tendency that increasingly transforms control over data into data "propertisation".

These developments result from a combination of legal mechanisms, including new legislative initiatives within established intellectual property regimes, as well as the application of traditional property and contract doctrines, to emerging technological realities. Privacy frameworks have also been developed to address specific concerns relating to personal data. Each of these frameworks pursues distinct objectives and operates through different doctrinal logics. Nevertheless, they converge in practice by enabling data holders to exercise varying degrees of control over access to and reuse of their data. Such control may arise indirectly from the protection of intangibles, from rights in physical infrastructure, from private ordering or from regulatory constraints grounded in personal-data protection. Collectively, these mechanisms contribute to what amounts to

[23] European Commission, "Data Market Study", 68–69, 82–83.

[24] OECD, "Mapping Relevant Data Collection Mechanisms for AI Training", 16, https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/10/mapping-relevant-data-collection-mechanisms-for-ai-training_62921889/3264cd4c-en.pdf (last accessed 27 March 2026).

[25] Competition and Markets Authority, "AI Foundation Models: Initial Report", available at https://assets.publishing.service.gov.uk/media/65081d3aa41cc300145612c0/Full_report_.pdf (last accessed 5 November 2025).

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

a process of data propertisation, even in the absence of explicit property rights in data.

In addition to these traditional legal frameworks, since 2015 the EU has pursued an extensive regulatory agenda on data governance. This growing body of norms introduces further layers of regulation governing access to and control over data and increasingly constitutes what is commonly referred to as the emerging "law of data". While these instruments reflect diverse regulatory ambitions – ranging from promoting innovation to safeguarding fundamental rights – they also add to an already fragmented legal landscape. The result is a patchwork of overlapping regimes that operate cumulatively to shape exclusionary outcomes, often without a coherent normative foundation.

This conceptual framing provides the foundation for the normative argument developed in the subsequent sections and underscores the need for a principled framework to assess when, if ever, exclusive rights in data are justified.

A. Traditional Frameworks: From Intellectual Property, through Possession, to Privacy Laws

The analysis that follows traces the different pathways through which data become enclosed across intellectual property, property law, private ordering and data protection. By mapping these mechanisms together, this section seeks to demonstrate how various forms of control, spanning from formalised legal frameworks to factual contingencies and exhibiting varying degrees of explicitness, are progressively converted into de facto exclusivity over data.

1. Intellectual property: copyright and sui generis database right

Intellectual property regimes provide a paradigmatic example of how legal protection designed for intangible subject matters may indirectly operate to restrict access to and use of data. Although copyright law does not protect data as such, it plays a central role in the broader process of data propertisation by extending exclusionary control over informational resources embedded in protected works and databases.

Copyright law protects intangible original works of authorship. Traditionally, it protects expressions, as opposed to ideas, facts, raw data or information as such.[26] In response to technological developments, the EU has adopted various legislative initiatives, two of which are particularly illustrative of the evolving regulation of control over and

[26] Berne Convention for the Protection of Literary and Artistic Works of 1886 (adopted 9 September 1886; amended 28 September 1979), art. 2; Agreement on Trade-Related Aspects of Intellectual Property Rights 1994 (adopted 15 April 1994), art. 9 ("TRIPS").

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

access to data. These initiatives mark, respectively, the early treatment of data in the context of digital technologies in the late 1990s and the most recent wave of extensive reform in the late 2010s, which already extends to AI technologies.

The first example is the EU Database Directive 1996, which harmonised legal protection for databases across Europe.[27] The Directive recognises two distinct forms of protection: copyright for databases demonstrating originality[28] and a *sui generis* right for databases involving "substantial investment" in obtaining, verifying or presenting their contents.[29] Consistently with copyright principles, copyright protection applies to the structure of databases and does not extend to the data they contain.[30] The *sui generis* right was introduced to address disparities in the protection of databases among Member States[31] and restricts the extraction or reutilisation of substantial parts of a database for 15 years.[32] The Directive limits rights holders' control by allowing "lawful users" to extract or reuse insubstantial parts[33] and enabling Member States to set specific exceptions.[34]

The *sui generis* right has faced significant criticism for restricting data access and stifling innovation.[35] Although Recital 46 clarifies that this right should not establish ownership over data per se, its practical effect often blurs the boundary between databases and the data they contain.[36] Therefore, legal protection directed at database structures and investment may function in practice as a gateway to controlling downstream data use. For example, the Court of Justice of the European Union ("CJEU") has broadly interpreted the scope of protection as precluding the extraction and reutilisation of certain geographical information taken from the claimant's geographical map, where such extraction is assessed qualitatively by reference to the investment reflected in the data and with particular regard to its potential commercial exploitation by third parties.[37] Under this approach, even an individual piece of information may fall within the scope of protection.[38] The European Commission's 2005 and 2018 evaluation reports found that, despite its intent, the right "comes close to protecting data as property".[39]

[27] European Parliament and Council Directive (EC) 9/1996 (OJ 1996 L 77 p.20) ("The Database Directive").
[28] Ibid., art. 3(1).
[29] Ibid., art. 7(1).
[30] Ibid., art. 3(2).
[31] Ibid., Recitals 1, 4.
[32] Ibid., art. 7(1), art. 10(1), Recitals 8, 39, 41.
[33] Ibid., art. 8(1).
[34] Ibid., art. 9(b).
[35] Hugenholtz, "Against 'Data Property'", 50.
[36] I. Stepanov, "Introducing a Property Right over Data in the EU: The Data Producer's Right – An Evaluation" (2020) 34 International Review of Law, Computers & Technology 65, 71.
[37] Case C-490/14, *Freistaat Bayern v Verlag Esterbauer GmbH* [2015] Bus. L.R. 1428.
[38] T. Synodinou, "Database Producer Protection: Between Rights and Liabilities" in T. Aplin (ed.), *Research Handbook on Intellectual Property and Digital Technologies* (Cheltenham 2020), 86.
[39] European Commission, "DG Internal Market and Services Working Paper: First Evaluation of Directive 96/9/EC on the Legal Protection of Databases", 24, available at https://digital-strategy.ec.europa.eu/en/library/staff-working-document-and-executive-summary-evaluation-directive-969ec-legal-protection-

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

The second example is a legislative initiative adopted almost two decades later: the Copyright in the Digital Single Market Directive 2019 ("Copyright Digital Market Directive").[40] Aimed at harmonising aspects of EU copyright law for digital and cross-border uses, the Copyright Digital Market Directive introduced a framework governing text and data mining – computer-based analysis of data to generate new knowledge. Text and data mining have become central to research and a key driver of innovation.[41] Because text and data mining may involve acts of reproduction, it can fall within the scope of copyright exclusivity.[42] Articles 3 and 4 of the Copyright Digital Market Directive define when such acts are permitted. Article 3 establishes a mandatory exception for research organisations and cultural heritage institutions with lawful access to works, for scientific research purposes.[43] Article 4, by contrast, applies to all users and allows rights holders to reserve their rights,[44] for example through technological measures or website terms of use.[45] Although the reservation must be proactive, namely rights holders must opt out of the exception, it also places a burden on users to verify the conditions of use and obtain permission where necessary.[46]

These provisions have sparked debate. Doctrinally, it was questioned whether they apply to AI training using copyright-protected works.[47] From a policy perspective, critics argue that these provisions strengthen exclusivity over protected works and databases and may hinder innovation.[48] Although copyright does not protect data as such, the

databases (last accessed 27 March 2026); European Commission, "Study in Support of the Evaluation of Directive 96/9/EC on the Legal Protection of Databases", 27, available at https://publications.europa.eu/resource/cellar/2d1f5a77-5982-11e8-ab41-01aa75ed71a1.0001.01/DOC_1 (last accessed 27 March 2026).

[40] European Parliament and Council Directive (EU) 790/2019 (OJ 2019 L 130 p.92) ("Copyright Digital Market Directive").

[41] C. Geiger, G. Frosio and O. Bulayenko, "Text and Data Mining: Articles 3 and 4 of the Directive 2019/790/EU", available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3470653 (last accessed 27 March 2026).

[42] A. Peukert, "Copyright in the Artificial Intelligence Act: A Primer" (2024) 73 GRUR International 497, 510.

[43] Copyright Digital Market Directive, art. 3, Recital 14; M.R.F. Senftleben, "Study on EU Copyright and Related Rights and Access to and Reuse of Data", 36–39, available at https://www.ivir.nl/publicaties/download/KI0822205ENN.en_.pdf (last accessed 27 March 2026).

[44] Copyright Digital Market Directive, art. 4; Senftleben, "Study on EU Copyright", 40.

[45] European Union Intellectual Property Office (EUIPO), "The Development of Generative Artificial Intelligence from a Copyright Perspective", 164–208, available at https://euipo.europa.eu/tunnel-web/secure/webdav/guest/document_library/observatory/documents/reports/2025_GenAI_from_copyright_perspective/2025_GenAI_from_copyright_perspective_FullR_en.pdf (last accessed 27 March 2026).

[46] N. Lucchi, "Generative AI and Copyright: Training, Creation, Regulation", 23, 26–27, available at https://www.europarl.europa.eu/RegData/etudes/STUD/2025/774095/IUST_STU(2025)774095_EN.pdf (last accessed 27 March 2026).

[47] P.G. Picht and F. Thouvenin, "AI and IP: Theory to Policy and Back Again: Policy and Research Recommendations at the Intersection of Artificial Intelligence and Intellectual Property" (2023) 54 Intellectual Property and Competition Law 916, 928; EUIPO, "Development of Generative Artificial Intelligence", 55.

[48] Geiger, Frosio and Bulayenko, "Text and Data Mining"; R. Ducato and A. Strowel, "Limitations to Text and Data Mining and Consumer Empowerment: Making the Case for a Right to 'Machine Legibility'" (2019) 50 International Review of Intellectual Property and Competition Law 649.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

ability of rights holders to opt out of text and data mining may nonetheless operate indirectly, in practice, as another form of de facto control over data. However, others view these provisions as a balanced compromise between proprietary control and access.[49]

Taken together, these copyright-based mechanisms illustrate how regimes formally detached from data ownership nonetheless contribute to data enclosure by converting exclusivity over protected works and databases into practical limits on data access and reuse.

2. Intellectual property: trade secrets

Trade-secret law, traditionally classified within intellectual property, illustrates a robust pathway to data exclusivity by focusing on information itself, rather than creative expression or organisational structure, based on its secrecy and commercial value. While formally framed as protection against unlawful acquisition, it operates in practice as a mechanism for converting informational secrecy into legally supported control over data.

According to the EU Trade-Secrets Directive, a trade secret means information that meets three requirements: it is secret in the sense that it is not generally known or accessible; it confers an economic benefit on its holder because the information is not publicly known; and it has been subject to reasonable steps to keep it secret.[50] The protection is provided only when the information at stake has commercial value that stems from its secrecy, that is, its commercial value would lapse by its disclosure.[51] Such a requirement means that the economic value of a trade secret is derived from keeping it secret and not necessarily from the mere investment in collecting the information.[52]

The protection provided for trade secrets is against the unlawful acquisition, such as by unauthorised access. Trade-secret protection provides limited exclusivity because various means for accessing the information are regarded as lawful. For example, "independent discovery" and "any other practice which [...] is in conformity with honest commercial practices" are regarded as legitimate ways of acquiring the information.[53] In other words, trade-secret protection does not treat the information as protected per se, but rather, it treats as unlawful certain ways for gaining access to the information. The

[49] A. Guadamuz, "A Scanner Darkly: Copyright Liability and Exceptions in Artificial Intelligence Inputs and Outputs" (2024) 73 GRUR International 111, 129.
[50] European Parliament and Council Directive (EU) 943/2016 (OJ 2016 L 157 p.1), art. 2 ("EU Trade-Secrets Directive").
[51] Ibid., art. 2(1)(b); TRIPS, art. 39(2)(b).
[52] S.K. Sandeen and T. Aplin, "Trade Secrecy, Factual Secrecy and the Hype Surrounding AI" in R. Abbott (ed.), *Research Handbook on Intellectual Property and Artificial Intelligence* (Cheltenham 2022), 456.
[53] EU Trade-Secrets Directive, arts. 3, 4(2)(a), 4(2)(b).

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

question of whether trade-secret and confidential-information protection is a property right is not new.[54] The courts occasionally use the notion of property with regard to information metaphorically to convey the idea of "ownership" over the information.[55]

Information that may be protected as a trade secret includes data and databases stored in digital form.[56] Trade-secret protection therefore constitutes another significant legal framework for exercising *control* over data, regardless of the absence of property rights in information as such and irrespective of whether trade secrets are classified as a "strong" form of property right or as a "quasi property", within a softer framework.[57] By grounding exclusivity in the secrecy and commercial value of information, including data, trade secrets provide a powerful mechanism for enclosing data, which may in practice significantly contribute to the cumulative process of data propertisation.

3. Possession

Outside the realm of intellectual property, property law doctrines governing tangible chattels provide another pathway through which control over data may be established. A particularly significant concept in this context is that of "possession" – a deeply rooted and powerful concept in property law,[58] under which the holding of a tangible "thing" or "object", which generates exclusive control over such "thing", may be recognised as legal ownership.[59] Thus, possession is a central element in establishing property rights by granting legal force to a de facto reality of exclusive control.[60] This legal move raises the fundamental question of whether possession is a fact or a right, as stating that someone "has possession" tends to convey, indirectly or by implication, the understanding that the law will recognise certain advantages akin to legal ownership.[61] The concept of possession originates in Roman law and has been adopted in both common law and continental legal systems.[62]

The concept of possession affects control over data in several ways. First, although data are intangible and non-rival, they are stored in tangible

[54] L. Bently, "Trade Secrets: 'Intellectual Property' But Not 'Property'?" in H. Howe and J. Griffiths (eds.), *Concepts of Property in Intellectual Property Law* (Cambridge 2013).
[55] Sandeen and Aplin, "Trade Secrecy", 182.
[56] T. Aplin et al., "The Role of EU Trade Secrets Law in the Data Economy: An Empirical Analysis" (2023) 54 International Review of Intellectual Property and Competition Law 826, 832–34.
[57] Ibid., at 832–35.
[58] S. van Erp, "Ownership of Data: The Numerus Clausus of Legal Objects" (2017) 6 Brigham-Kanner Property Rights Conference Journal 235.
[59] L. Rostill, *Possession, Relative Title, and Ownership in English Law* (Oxford 2021), 1–2.
[60] Ibid., at 15; S. van Erp and B. Akkermans, *Cases, Materials and Text on Property Law* (Oxford 2012), 99, 193, 206–08.
[61] van Erp, "Ownership of Data", 235–36.
[62] D. Lueck, "The Rule of First Possession and the Design of the Law" (1995) 38 Journal of Law & Economics 393, 394.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

devices such as computers and servers.[63] As physical objects, servers are subject to exclusive holding and control.[64] Thus, the entity that collects and stores data gains possession of the devices storing the data (i.e. the servers), thereby effectively controlling the data stored and the ability to limit access to them. The physical possession of tangible objects thus indirectly extends control to the data they contain, generating an additional, significant mechanism of de facto data propertisation.[65]

The path of control over data through physical possession of the tangible data container may raise practical concerns. For example, a data collector who stores data using the services of a third party, such as an external cloud provider, does not exercise direct physical control over the storage infrastructure. The question then arises whether contractual arrangements, together with the application of the concept of "possession through another", recognised in traditional property law (e.g. where physical control over a tangible asset is exercised on behalf of another person),[66] may be extended to the context of data storage, thereby maintaining control over the data.[67]

It should be noted that physical possession of a tangible data container may also be subject to legally recognised exclusivity under the trade-secrets framework.[68] However, for information, that is, data, to qualify for trade-secret protection, certain requirements must be met, including that the information is kept secret and derives independent commercial or economic value from its secrecy.[69] Physical possession of a tangible data container, which may indirectly generate a form of legally recognised exclusivity over the data itself, should therefore be distinguished from intellectual property frameworks such as trade-secret protection, which are subject to stricter eligibility requirements for conferring exclusivity over information.

A second, and more challenging, path through which the concept of possession may affect control over data concerns its application directly to non-tangible data itself. The concept of possession, as developed in traditional property law, is rooted in tangible assets.[70] By contrast, in intellectual property law, which grants exclusive rights over intangible subject matter, the concept of possession is more complex[71] and is often

[63] van Erp, "Ownership of Data", 250–51.
[64] Zech, "Information as Property", 194–95.
[65] M. Eckardt and W. Kerber, "Property Rights Theory, Bundles of Rights on IoT Data, and the EU Data Act" (2024) 57 European Journal of Law and Economics 113, 118–19.
[66] Rostill, *Possession*, 21–22.
[67] J.D. Michels and C. Millard, "The New Things: Property Rights in Digital Files?" [2022] C.L.J. 323, 323–24.
[68] J.H. Reichman, "How Trade Secrecy Law Generates a Natural Semicommons of Innovative Know-How" in R.C. Dreyfuss and K.J. Strandburg (eds.), *The Law and Theory of Trade Secrecy: A Handbook of Contemporary Research* (Cheltenham 2011), 196.
[69] EU Trade-Secrets Directive, art. 2.
[70] L. Rostill, "Relative Title and Deemed Ownership in English Personal Property Law" (2015) 35 O.J.L.S. 31.
[71] See e.g. the discourse regarding the expansion of the tort of conversion to intangibles; S. Green, "To Have and to Hold? Conversion and Intangible Property" (2008) 71 M.L.R. 114.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

understood as having evolved metaphorically[72] or as a by-product of legal exclusivity.[73] Many doctrines in intellectual property law may have some linkage to the notion of possession, based on the notion of "first holding",[74] albeit often by way of analogy.[75] The question, then, is whether data ownership has evolved into a distinct and legally recognised concept,[76] based on similar metaphorical foundations, such as the notion of "first holding" that is found in intellectual property law. It has been suggested that the idea of data ownership represents a broader reconceptualisation of "ownership" as "control",[77] which, though closely related, remains conceptually distinct.[78] In the ongoing evolution of legal frameworks governing data, such a development may represent a future trajectory, signalling a significant paradigm shift towards the direct propertisation of data, with far-reaching implications for the data economy.[79]

4. Private ordering: contractual arrangements and technological measures

Beyond formal proprietary regimes, private ordering provides another powerful pathway to establishing control and thereby exclusivity over data. Through contracts and technological measures, data holders may transform factual control into enforceable restrictions on access and reuse, even in the absence of established property rights in data.

Contracts play a crucial role in securing commercial interests associated with data (e.g. data sharing agreements, data service agreements).[80] Through contractual arrangements across a wide range of relationships, data holders may acquire exclusive control, which, although effective only *inter partes*, may, when applied extensively across multiple contractual relationships, such as through standard-form contracts, generate broad de facto control over data.[81] The use of private ordering through contractual means has been addressed in copyright law, particularly in relation to the question of whether exceptions and limitations should be treated as mandatory,

[72] S. van Erp, "The Future of Private Law between Sustainability Challenges and the Digital Revolution: Managed Retreat and Digital Assets as Examples" in L. Antoniolli and P. Iamiceli (eds.), *The Making of European Private Law: Changes and Challenges* (Trento 2024), 30.
[73] van Erp, "Ownership of Data", 242–43.
[74] D. Oliar and J.Y. Stern, "Right on Time: First Possession in Property and Intellectual Property" (2019) 99 Boston University Law Review 395.
[75] P. Mysoor, "Possession in Copyright by Analogy" (2024) 140 L.Q.R. 277.
[76] Michels and Millard, "New Things", 338–39; J. Verstappen, "Data Ownership" in S. van Erp and K. Zimmermann (eds.), *Research Handbook on European Property Law* (Cheltenham 2024).
[77] van Erp, "Future of Private Law", 29–31.
[78] van Erp, "Ownership of Data", 244, 251–52.
[79] Michels and Millard, "New Things", 351–52.
[80] T. Margoni and A.M. Strowel, "Contractual Freedom and Fairness in EU Data Sharing Agreements" in J. de Werra and I. Calboli (eds.), *Research Handbook on Intellectual Property Licensing* (Cheltenham 2025).
[81] L. Guibault, "Wrapping Information in Contract: How Does it Affect the Public Domain?" in L. Guibault and P.B. Hugenholtz (eds.), *The Future of the Public Domain* (Alphen aan den Rijn 2006), 97–98.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

even in the absence of an explicit statutory provision or whether they may be circumvented through the exercise of contractual freedom, thereby effectively expanding the scope of exclusivity.[82] Similar effects may be observed in relation to the use of contractual arrangements to exercise control over unprotected subject matter, including data.[83]

Technological measures, including encryption and access controls, are commonly used safeguards to protect data from unauthorised access and breaches of contractual obligations.[84] These technological measures may grant actual control over access to data and its use in various circumstances, beyond legal exclusivity. A prominent example is technological protection measures used in relation to copyright-protected works, which can prevent access not only to the copyright-protected content but indirectly also to unprotected data included within that content.[85] While the CJEU determined that the use of such measures is legitimate insofar as they are proportionate and closely tied to the prevention of copyright infringement,[86] the application of these standards nonetheless leaves rights holders considerable leeway, in practice, to exercise extensive control.[87]

Therefore, private ordering frameworks, such as contractual arrangements and technological measures, establish control over data, which generates yet another de facto exclusivity regime,[88] adding a significant element to the cumulative tendency of data propertisation.

5. Privacy and personal-data rights-related arguments

The complex legal frameworks addressing data are further complicated by an additional prominent body of law introducing distinct concerns relating to privacy and personal data. This regime regulates the collection, use and control of personal information in a comprehensive and systematic manner.

Personal-data protection is governed in the EU by the General Data Protection Regulation ("GDPR"), which establishes strict requirements for collecting, using and storing personal data by entities operating within the EU or offering goods or services to EU residents.[89] The GDPR aims to grant individuals greater transparency and control over the processing of their personal information, including the right to

[82] M. Kretschmer et al., "The Relationship between Copyright and Contract Law", 89–91, 105–06, available at https://webarchive.nationalarchives.gov.uk/ukgwa/20140603083549/https://www.ipo.gov.uk/ipresearch-relation-201007.pdf(last accessed 27 March 2026).
[83] Guibault, "Wrapping Information", 97–98.
[84] D.J. Solove and W. Hartzog, *Breached! Why Data Security Law Fails and How to Improve It* (Oxford 2022), ch. 7.
[85] InfoSoc Directive, art. 6.
[86] Case C-355/12, *Nintendo Co. Ltd. v PC Box Srl.* [2014] E.C.D.R. 6.
[87] Moscon, "Data Access Rules", 10.
[88] Senftleben, "Study on EU Copyright", 26–27.
[89] European Parliament and Council Regulation (EU) 679/2016 (OJ 2016 L 119 p.1) ("GDPR").

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

access, correct and delete data pertaining to them. The GDPR also establishes rules regarding data portability.[90] A central issue addressed by the GDPR is the definition of personal data, framed broadly to include data relating to an identified or identifiable natural person.[91]

Yet, although personal-data protection regulation is designed to safeguard the interests of individual data subjects, data controllers may invoke it as a means of exercising control over data and avoiding data sharing for commercial reasons. In such cases, the asserted protection of personal information may function less as a genuine privacy safeguard and more as a strategic mechanism for maintaining control over data. For example, personal-data protection arguments may be deployed strategically to resist sharing data for competition-related reasons and to preserve market dominance, even when technical measures could effectively address privacy concerns.[92] In this way, privacy regulation may operate, in practice, as an implicit and indirect exclusionary framework, providing data holders with an additional vehicle for extracting commercial value from their mere control over data.[93] While personal-data protection rules are therefore relevant to the discussion of data control, they will not be further addressed in detail.

* * *

In sum, the key frameworks reviewed reveal various mechanisms through which greater control over data is achieved, even if only indirectly. Taken together, these mechanisms depict a cumulative trajectory towards data propertisation. This cumulative effect should not be overlooked, as the growing number of regimes conferring actual control over data may ultimately amount to a substantive transformation in the perception of data as an asset subject to exclusivity.

B. The New EU Data Regulation: Transitions between Exclusivity and Pro-Access Approaches to Data

The increasingly expansive mechanisms addressing control over data within traditional legal frameworks have been accompanied by a further major regulatory development: the emergence of EU data law. Since 2015, the EU has been promoting an extensive body of legislation on data governance,[94] reflecting a fundamental tension between competing approaches to data regulation. While some regulatory initiatives adopt a pro-access orientation, promoting a non-proprietary approach to data,

[90] Ibid., art. 20.
[91] Ibid., art. 4(1). This broad definition is addressed in Digital Omnibus Proposal, 54–55.
[92] Crémer, de Montjoye and Schweitzer, "Competition Policy", 77–78.
[93] H. Zech, "A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data" (2016) Journal of Intellectual Property Law & Practice 4.
[94] European Commission, "Digital Single Market Strategy".

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

others preserve control and exclusivity over data, directly or indirectly, thereby reinforcing the tendency towards data propertisation. EU data law is a highly dynamic and ongoing project, responding to multiple and often conflicting considerations, generating a complex web of regulatory frameworks addressing data.[95]

The first EU milestone clearly reflecting a pro-access and non-proprietary approach to data was Directive (EC) 98/2003 ("PSI Directive"),[96] later amended and renamed by Directive 1024/2019 ("Open Data Directive").[97] Its objective is to increase the availability of public-sector information across the EU. The PSI Directive requires public bodies to make information available for reuse by electronic means, where possible and appropriate, subject only to limited fees.[98] The Open Data Directive builds on this foundation by promoting real-time access to data through appropriate technical means and expanding the categories of public data available for reuse, including from public undertakings and research organisations.[99] It also seeks to prevent exclusive arrangements that could limit data accessibility. In essence, the Open Data Directive broadens the scope of public-sector data that should remain open to the public.[100] Alongside the PSI Directive, these instruments advance the principle that data generated and held by public bodies should be publicly accessible.[101]

Aligned with these frameworks, the EU's Data Governance Act aims to establish a single European data space,[102] enabling the secure sharing of certain public-sector data among businesses, public bodies and individuals.[103] While the Act reflects a pro-access approach, its practical effect remains limited, as it applies only to specific categories of public data already addressed by the PSI and Open Data Directives.

While the PSI and Open Data Directives promoted broad access to public-sector information, a more cautious stance emerged with the adoption, in 2022, the Digital Services Act ("DSA").[104] Although primarily aimed at ensuring a safe and accountable online environment, the DSA also reflects a more limited pro-access data policy, particularly regarding private-sector data. Article 40 DSA regulates access to data collected and

[95] Digital Omnibus Proposal.
[96] European Parliament and Council Directive (EC) 98/2003 (OJ 2003 L 345 p.1) ("PSI Directive").
[97] European Parliament and Council Directive (EU) 1024/2019 (OJ 2019 L 172 p.56) ("Open Data Directive").
[98] PSI Directive, arts. 3, 4, 6.
[99] Ibid., Recital 4.
[100] M. van Eechoud, "Study on the Open Data Directive, Data Governance and Data Act and Their Possible Impact on Research", available at https://www.ivir.nl/publicaties/download/KI0822204ENN.en_.pdf (last accessed 24 April 2026).
[101] Open Data Directive, art. 12.
[102] European Parliament and Council Regulation (EU) 868/2022 (OJ 2022 L 152 p.1), Recital 2 ("Data Governance Act").
[103] Ibid., art. 3.
[104] European Parliament and Council Regulation (EU) 2065/2022 (OJ 2022 L 277 p.1) ("DSA").

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

stored by online platforms, subject to oversight by a designated coordinator. These data are valuable for research across sectors, yet access is restricted to digital service coordinators and vetted researchers for the specific purpose of studying systemic risks in the EU.[105] Therefore, although Article 40 grants some researchers limited access to previously undisclosed data, enabling new research opportunities, it does not open data to the broader public and its strict limitations raise doubts about its effectiveness. Ultimately, the DSA adopts a restrained approach to data sharing, leaving most platform-held data under an exclusive control regime. Public criticism of these limitations prompted the European Commission to launch a public consultation[106] and, in response, various opinions called for greater access to these data.[107]

It may be argued that the difference in approaches between the Open Data Directive and the DSA stems from the distinction between data originating from the public and the private sectors. Yet, in 2023 the EU adopted the Data Act,[108] which reflects an attempt to foster an open-access approach to private-sector data through the establishment of new data access and data-sharing rights for users.[109] Therefore, while the EU Data Act constitutes a milestone in the evolving project of EU data law, its conceptual turn highlights the inconsistency in the EU's approach to private-sector data.

The EU Data Act sets rules for the use of data generated by connected products, defined as items that generate or collect data about their use or their environment, as well as by services related to those products.[110] It requires that such data be made "by default, easily, securely, free of charge, in a comprehensive, structured, commonly used and machine-readable format"[111] and grants users the right to share data with third parties, subject to specific conditions.[112] Internet of Things ("IoT") devices are connected products supported by related digital services and are therefore notably regulated under the Act.[113] In this way, the Act

[105] Ibid., art. 40(4)–(11).
[106] European Commission, "Call for Evidence for an Initiative: Delegated Regulation on Data Access Provided for in the Digital Services Act", available at https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13817-Delegated-Regulation-on-data-access-provided-for-in-the-Digital-Services-Act_en (last accessed 24 April 2026).
[107] P. Leerssen, "Call for Evidence on the Delegated Regulation on Data Access Provided for in the Digital Services Act: Summary and Analysis", available at: https://digital-strategy.ec.europa.eu/en/library/digital-services-act-summary-report-call-evidence-delegated-regulation-data-access (last accessed 24 April 2026).
[108] European Parliament and Council Regulation (EU) 2854/2023 (OJ 2023 L 22 p.1) ("EU Data Act").
[109] M. Leistner and L. Antoine, "Attention, Here Comes the EU Data Act! A Critical In-Depth Analysis of the Commission's 2022 Proposal" (2022) 13 Journal of Intellectual Property, Information Technology and E-Commerce Law 339, 340.
[110] EU Data Act, art. 2.
[111] Ibid., art. 3.
[112] Ibid., art. 5.
[113] Ibid., Recital 1.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

promotes data sharing by private-sector entities, offering potential benefits for consumers, businesses and society at large.[114]

By rejecting attempts to adopt a new *sui generis* data producers' right[115] and adhering instead to a pro-data-sharing approach, the EU Data Act reflects a turning point in the EU's treatment of data, extending the pro-access agenda beyond public-sector data to data held by private actors. This move is of great conceptual importance and marks a significant leap in the evolution of EU data law,[116] nevertheless the EU Data Act has been criticised on various grounds. A major concern relates to significant uncertainties surrounding the practical implementation of the data access and data-sharing mechanisms,[117] as well as the Act's relationship with other legal frameworks, such as copyright and the *sui generis* database right[118] and other data regulations.[119] One key uncertainty involves coordination with trade-secrets protection, particularly regarding the implementation of the exception in the EU Data Act that allows a data holder to oppose data sharing if it may cause severe economic harm to trade secrets.[120] Moreover, the EU Data Act's impact is limited by its narrowed definition of connected products, which excludes computers and mobile phones,[121] and instead covers devices such as household appliances or industrial and agricultural machinery. Furthermore, broader concerns have been raised that the EU Data Act may fail to fulfil its objectives because of various market failures, including flaws in the data-sharing mechanism and barriers raised by the de facto control of data holders in the private sector.[122]

Alongside the EU Data Act, another instrument supporting greater access to privately held data is the Digital Markets Act ("DMA"), adopted in 2022.[123] The DMA seeks to make EU digital markets fairer and more contestable by imposing obligations on large digital platforms designated as "gatekeepers".[124] Some of these obligations concern data access – most notably, requiring gatekeepers to grant business users access to data generated through their use of platform services, subject to certain

[114] Ibid., Recitals 1, 2; F. Casolari, C. Buttaboni and L. Floridi, "The EU Data Act in Context: A Legal Assessment" (2023) 31 International Journal of Law and Information Technology 399, 406.
[115] Moscon, "Data Access Rules", 13.
[116] Casolari, Buttaboni and Floridi, "EU Data Act", 408.
[117] Eckardt and Kerber, "Property Rights Theory", 115.
[118] Although Article 43 of the EU Data Act addresses the relationship with the *sui generis* database right, it leaves many uncertainties unresolved. J. Drexl et al., "Position Statement of the Max Planck Institute for Innovation and Competition on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act)", 96–100, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4136484&download=yes (last accessed 24 April 2026).
[119] Ibid., at 5; Casolari, Buttaboni and Floridi, "EU Data Act", 409.
[120] EU Data Act, arts. 4(3)(b), 5(8)(b); Drexl et al., "Position Statement of 23 February 2022", 100–05.
[121] EU Data Act, Recital 15.
[122] Drexl et al., "Position Statement of 23 February 2022", 16.
[123] European Parliament and Council Regulation (EU) 1925/2022 (OJ 2022 L 265 p.1).
[124] Ibid., Recital 7, art. 3.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

limitations.[125] This reflects the view that gatekeepers' control over data entrenches market power, which should be mitigated through greater data access to foster competition.[126]

The DMA also enhances end users' access to their own generated data,[127] reinforcing individual control consistent with the GDPR's data portability provisions.[128] However, unlike the EU Data Act, its data access rights apply only to gatekeepers – dominant platforms such as Alphabet and Meta.[129] The DMA thus mandates private-sector data sharing based on competition law principles.[130]

As with the EU Data Act, questions persist regarding the practical implementation of data access and portability, including technological and legal barriers.[131] Moreover, as a competition-oriented regulation, the DMA has faced criticism for uncertainties in its relationship with existing competition law frameworks[132] and doubts about its effectiveness in lowering market entry barriers.[133]

To sum up, the evolving EU project of data law is adding a substantial regulatory layer addressing questions of control over and access to data, alongside traditional legal mechanisms. However, the EU has not assembled the various pieces of legislation under one umbrella and the treatment of these questions remains fragmented. This policy has been criticised because the result is a legislative labyrinth of occasionally inconsistent provisions, ultimately making it difficult to understand the applicable law.[134] The regulatory approach to actual control over privately held data is particularly inconsistent and the resulting uncertainty could have a chilling effect on innovation.[135]

IV. The Need to Shift from Exclusivity to an Open-Access Regime

The legal landscape described above, whether based on traditional legal frameworks or on existing EU data regulation, illustrates the conflict

[125] Ibid., art. 6(10).
[126] Ibid., Recital 46; P.G. Picht, "Caught in the Acts: Framing Mandatory Data Access Transactions under the Data Act, Further EU Digital Regulation Acts, and Competition Law" (2023) 14 Journal of European Competition Law & Practice 67, 69.
[127] DMA, art. 6(9).
[128] J. Krämer, "Data Access Provisions in the DMA", 6, available at https://cerre.eu/wp-content/uploads/2022/11/DMA_DataAccessProvisions-2.pdf (last accessed 24 April 2026).
[129] Schweitzer et al., "Data Access", 338, 343.
[130] Crémer, de Montjoye and Schweitzer, "Competition Policy", 91–92.
[131] Schweitzer et al., "Data Access", 340.
[132] J. Drexl et al. "Position Statement of the Max Planck Institute for Innovation and Competition of 2 May 2023 on the Implementation of the Digital Markets Act (DMA)" (2023) 72 GRUR International 864, 867–68.
[133] S. Unekbas, "Verticalization of Data Sharing and the Difficult Path to 'Eunnovation'" (2023) 7 Market and Competition Law Review 131.
[134] T. Streinz, "The Evolution of European Data Law" in P. Craig and G. de Búrca (eds.), *The Evolution of EU Law* (Oxford 2021), 902–03; Schweitzer et al., "Data Access", 340.
[135] Draghi, "European Competitiveness – Part A", 30.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

between mechanisms that facilitate legal protection of data and the competing need to enable data access, sharing and use for fostering innovation, particularly in AI technologies. The tendency to expand control over data evokes the enduring tension in intellectual property law between natural law theories and utilitarian or instrumental approaches.

Natural law justifications for intellectual property rights are traditionally associated with labour and personality theories and are most commonly invoked in relation to creative endeavours.[136] In the context of data, natural law approaches are particularly evident in the concept of possession, which transforms the fact of physical holding – a "natural" situation – into a legal entitlement.[137] Yet the question remains why such legal recognition, even if only indirect, should apply to data, an "industrial" intangible that does not involve creative or personality-based endeavours.[138] This question is particularly pressing when data are collected without substantial labour.

Utilitarian approaches, by contrast, focus on legal entitlements that maximise public benefit and are closely tied to economic reasoning.[139] Since EU policy-makers explicitly frame data governance as a means to promote the data economy, it should primarily rest on utilitarian and economic principles.[140]

Focusing on utilitarian approaches highlights the economic and social dilemmas inherent in developing a coherent framework for information regulation: the need to incentivise the development of databases and information products, while also ensuring access to raw data.[141] This issue, debated since the 1990s, has been reignited by the rise of AI and the growing centrality of data, bringing it to the forefront of contemporary legal discourse.[142] In today's environment, where the data economy drives growth and welfare, policy-makers – particularly in the EU – seek to promote data-driven innovation.[143] To this end, policy instruments aim to enhance the efficient use of data through mechanisms that foster sharing.[144] Within this discourse, traditional utilitarianism and law-and-economics approaches both appear,[145] yet it seems necessary to revisit the incentive theory and its assumption of market failure as applied to data.[146]

[136] Drahos, *Philosophy*, 47, 82.
[137] See Part III(A)(3) above; van Erp, "Ownership of Data", 235–36.
[138] Zech, "Information as Property", 194–95.
[139] Drahos, *Philosophy*, 251.
[140] European Commission, "Digital Single Market Strategy", 2; European Commission, "Shaping Europe's Digital Future", 2–3.
[141] Ullrich, "Intellectual Property", 4, 20–21, 27.
[142] OECD, "Digital Economy Outlook", 29–31; Draghi, "European Competitiveness – Part A", 30.
[143] European Commission, "Shaping Europe's Digital Future", 2–3.
[144] Ibid., at 7; EU Data Act, Recitals 2, 4.
[145] Eckardt and Kerber, "Property Rights Theory", 117; Crémer, de Montjoye and Schweitzer, "Competition Policy", 90.
[146] Zech, "Information as Property", 194.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

According to incentive theory, proprietary protection for intangibles facilitates investment recovery and profit generation that serve the public interest in innovation.[147] Propertisation thus enhances market efficiency by addressing a market failure, that is, the lack of incentives to innovate.[148] Traditional law-and-economics theories further view property rights as promoting efficient resource allocation (e.g. the Coase theorem).[149] Yet, in an era of large-scale data accumulation at modest cost, it is doubtful that a market failure exists requiring additional incentives for data collection and storage.[150] As the cost of data-related activities falls, the need for incentives declines.[151] Recognising property rights may paradoxically hinder innovation, since exclusivity restricts public use of intangibles, creating a social cost without a corresponding social benefit. Property rights in data may therefore impede data-driven innovation, undermining efficiency and societal welfare.[152]

Similar criticism may be raised with respect to a further strand of utilitarian justifications, which focus not on incentivising the *ex-ante* creation of intangibles, but on their efficient use after creation, for example through refinement and further development, and on preventing underutilisation.[153] This justification for extending control over data collection raises serious concerns. It risks entrenching anti-competitive effects that undermine utilitarian goals[154] and it is of doubtful relevance in the context of data, which function as the raw "materials" of data-driven markets and, particularly in the age of AI,[155] are not plausibly subject to risks of underutilisation.

Moreover, as investment becomes less decisive in determining whether to engage in data collection, the justification for addressing a potential market failure diminishes.[156] The validity of the incentive theory as a basis for data property thus erodes, as data collection, in many instances, is ancillary to the

[147] Drahos, *Philosophy*, 22–32.
[148] Ibid.
[149] R.H. Coase, "The Problem of Social Cost" (1960) 4 Journal of Law & Economics 386; Y. Barzel, *Economic Analysis of Property Rights*, 2nd ed. (Cambridge 1997).
[150] Kerber, "Specifying and Assigning"; H. Beverley-Smith, "Rights in Data and Information" in R. Dreyfuss and J. Pila (eds.), *The Handbook of Intellectual Property Law* (Oxford 2017), 597.
[151] E. Mathieu, "The Price of Computer Storage has Fallen Exponentially since the 1950s", available at https://ourworldindata.org/data-insights/the-price-of-computer-storage-has-fallen-exponentially-since-the-1950s (last accessed 25 April 2026).
[152] Eckardt and Kerber, "Property Rights Theory", 121; Schweitzer et al., "Data Access", 70–72, 119, 128. For a counter-argument supporting property rights in data, see e.g. A. Boerding et al., "Data Ownership: A Property Rights Approach from a European Perspective" (2018) 11 Journal of Civil Law Studies 323; F. Szilagyi, "The Necessity of Data Allocation: A Plea for a Private Law (Property Law) Perspective" (2021) 10 European Property Law Journal 180.
[153] M.A. Lemley, "Ex Ante Versus Ex Post Justifications for Intellectual Property" (2004) 71 University of Chicago Law Review 129.
[154] Ibid., at 131–32.
[155] OECD, "Mapping Relevant Data".
[156] Beverley-Smith, "Rights in Data", 597; Zech, "Information as Property", 194–95.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

primary and intentional activities of data collectors.[157] Much of the data now held by entities were collected as by-products of primary business operations that would have occurred regardless of legal protection. Therefore, one could argue that profits from such incidentally collected data are largely opportunistic, since their collection was not contingent on proprietary expectations.[158] This is particularly evident in the case of so-called "spin-off" databases, where data originally collected for a certain purpose is later repurposed or monetised in ways that go beyond the scope of the original activity,[159] raising further questions about the appropriateness of proprietary claims.

In the absence of a demonstrable need for proprietary rights to establish data collection and storage activities, especially given falling data-storage costs[160] and, without reliance on the potential profits from this activity per se, it is difficult to justify ownership claims over the results of such activities. Although data can yield commercial advantages to its holder, we argue that in the absence of a normative justification of property rights, not every commercial advantage created by collected data should result in exclusive rights.[161] The granting of such rights is the flaw in the logic underlying the proprietary tendency, which induces a legal property right from the fact that an entity performed an action on data and, therefore, claims ownership of them.[162] This circular reasoning has been often observed in the legal treatment of information.[163]

The model of data sharing and the denial of proprietary control may, in many cases, have a positive effect by fostering market development and competitiveness for the benefit of the public.[164] The public interest in maximising data accessibility may justify adopting a guiding principle of a default non-proprietary regime.[165] Only where substantial incentives are needed to collect and store data, and the result serves the public good, can limited exclusivity be justified.[166] Such a norm may also constrain the legal mechanisms that leverage possession into a de facto control

[157] J. Drexl, "Designing Competitive Markets for Industrial Data" (2017) 8 Journal of Intellectual Property, Information Technology and E-Commerce Law 257, 260.
[158] W. Kerber, "A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis" (2016) 65 GRUR International 989.
[159] E. Derclaye, "Databases 'Sui Generis' Right: Should We Adopt the Spin-Off Theory?" (2004) 26 European Intellectual Property Review 402.
[160] US Bureau of Labor Statistics, "Historical Price of Computer Memory and Storage", available at https://ourworldindata.org/grapher/historical-cost-of-computer-memory-and-storage (last accessed 5 November 2025).
[161] Hugenholtz, "Against 'Data Property'", 60.
[162] J. Waldron, "From Authors to Copiers: Individual Rights and Social Values in Intellectual Property" (1993) 68 Chicago-Kent Law Review 841.
[163] W.J. Gordon, "On Owning Information: Intellectual Property and the Restitutionary Impulse" (1992) 78 Virginia Law Review 149.
[164] Schweitzer et al., "Data Access", 71.
[165] S.K. Sandeen and U.-M. Mylly, "Trade Secrets and the Right to Information: A Comparative Analysis of EU and US Approaches to Freedom of Expression and Whistleblowing" (2020) 21 North Carolina Journal of Law & Technology 1.
[166] Kerber, "Governance of IoT Data".

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

over data, including more institutionalised legal frameworks such as trade-secrets protection[167] and private ordering through technological measures or contracts.[168] Recognising the circular reasoning that drives proprietary impulses and its adverse effects on innovation, particularly in AI, is a necessary step towards establishing a coherent and public-oriented data policy.

V. Data Taxonomy: What, Who and Why?

In the rapidly evolving landscape of data regulation, the importance of data taxonomy cannot be overlooked. Data taxonomy serves several critical functions in regulation: as a foundational method for categorising data to enhance their understanding;[169] to provide a common language for stakeholders in various sectors, facilitating clearer communication and understanding of data-related issues; and to ensure consistency of the increasingly complex regulatory requirements. The development of a robust taxonomy is essential for creating a legal system that is both comprehensive and flexible to accommodate the dynamic nature of data and can serve as a basis for designing coherent foundations for norms.[170] Last, a taxonomy makes possible the development of regulatory approaches that can be tailored to the unique characteristics of various data categories.

We will examine three primary questions: What are the data characteristics? Who collected the data? Why were the data collected?

A. What?

The "what" classification of data refers to their unique characteristics. One of the most critical distinctions is between "identifying" and "non-identifying data". "Identifying data" contain personal information that can be directly linked to an individual, such as names and biometric data.[171] These data are subject to personal-data regulations, which significantly limit their accessibility. "Non-identifying data", by contrast, have been stripped of all personally identifying information. They may include aggregated statistics or anonymised datasets that cannot be traced back to individuals. Although "non-identifying data" offer greater flexibility of usage and sharing, advanced data analysis techniques can, occasionally, re-identify individuals from anonymised data, raising

[167] Sandeen and Mylly, "Trade Secrets", 19.
[168] Drexl, "Designing".
[169] D.H. Doty and W.H. Glick, "Typologies as a Unique Form of Theory Building: Toward Improved Understanding and Modeling" (1994) 19 Academy of Management Review 230; K.K.D. Bailey, *Typologies and Taxonomies: An Introduction to Classification Techniques* (London 1994).
[170] R.C. Nickerson, U. Varshney and J. Muntermann, "A Method for Taxonomy Development and Its Application in Information Systems" (2013) 22 European Journal of Information Systems 336.
[171] GDPR, art. 4(1).

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

ongoing privacy concerns.[172] Consequently, the distinction between identifying and non-identifying data – which underpins the analytical matrix proposed in this paper – must be understood as ideal-typical, rather than absolute. Mixed datasets that include both identifying and non-identifying elements present particular challenges, especially when data separation entails significant costs or technical barriers.[173] Ensuring that data remains truly non-identifying may require not only technical safeguards but also continuous monitoring of re-identification risks, particularly as analytical capabilities evolve.

The level of data organisation and refinement is another key characteristic of data. Structured data are well organised, typically stored in databases or spreadsheets with clearly defined fields and relationships. This type of data is easily searchable and analysable using standard data processing tools and commonly reflects deliberate investment and commercial intent.[174] By contrast, raw or unstructured data often collected incidentally and lack a predefined format or organisation. This can include text documents, social media posts or sensor outputs.[175] Although raw data retain their original form and potentially valuable nuances, they often require significant processing and cleansing before they can be effectively analysed. Therefore, their less-processed nature may indicate the absence of purposeful commercial exploitation.

B. Who?

Understanding who collects and stores the data is crucial for assessing its reliability, accessibility and potential embedded biases. Data collection is carried out by both private and public entities, each operating under distinct motivations and regulatory frameworks.

Private entities, such as corporations, collect data to inform business decisions, improve products or services and gain competitive advantages.[176] Examples include e-commerce platforms collecting purchase history and social media companies tracking user interactions. Public entities, including government agencies and academic institutions, collect data for purposes such as policy-making, administrating public services, scientific research and ensuring public welfare.[177] Examples

[172] M. Altman et al., "Towards a Modern Approach to Privacy-Aware Government Data Releases" (2017) 30 Berkeley Technology Law Journal 1899.
[173] I. Graef, R. Gellert and M. Husovec, "Towards a Holistic Regulatory Approach for the European Data Economy: Why the Illusive Notion of Non-Personal Data Is Counterproductive to Data Innovation" (2019) 44 European Law Review 605.
[174] A. Jonker and A. Gomstyn, "Structured vs. Unstructured Data: What's the Difference?", available at https://www.ibm.com/think/topics/structured-vs-unstructured-data (last accessed 25 April 2026).
[175] J. Sedlakova et al., "Challenges and Best Practices for Digital Unstructured Data Enrichment in Health Research: A Systematic Narrative Review" (2023) 2 PLOS Digital Health.
[176] European Commission, "Shaping Europe's Digital Future", 2–3.
[177] K. Hardy and A. Maurushat, "Opening Up Government Data for Big Data Analysis and Public Benefit" (2017) 33 Computer Law & Security Review 30, 31.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

include census data collected by national statistics offices and health data gathered by public health departments.

The differentiation between public-sector and private-sector entities is often characterised by their institutional status as "for-profit" or "not-for-profit". However, this distinction is not entirely binary. The term "hybrid organisations" is used to describe entities that combine market-oriented activities with social or public-interest objectives. This category most commonly refers to for-profit entities engaged in socially driven activities, but it may also include not-for-profit entities that engage in commercially oriented activities.[178] All entities across this spectrum may be engaged in data collection.

C. Why?

The motivations behind data collection profoundly influence how data are gathered, used and shared. Data collection can be categorised based on the level of intentionality behind it. Intentional data collection can also be treated as value-driven; it occurs when an entity deliberately gathers data for a specific purpose.[179] This type of data collection gathers data with the explicit intent of deriving insights, making predictions or creating new products or services. This approach often involves strategic planning and significant investment in data analytics capabilities.[180]

Incidental, casual or passive data collection, also known as operational data collection, occurs as a by-product of other activities or systems, without the data themselves being the object of the activity or a primary asset at the time of collection.[181] This can include data generated through Internet browsing, mobile phone usage or operation of IoT devices, or as a result of transaction records or customer service interactions. Although incidental data collection can provide vast amounts of data, it may raise concerns regarding consent and privacy.[182]

Finally, the commercial intent behind data collection significantly influences the data collection objectives. Commercial data collection is driven primarily by profit motives, aiming to improve business operations, target marketing efforts or develop new revenue streams.[183] This type of data collection is subject to market forces and competitive pressures. Non-commercial data collection, by contrast, is typically aimed

[178] C. Mereghetti and M. Nuccio, "Towards a Taxonomy of Hybrid Organizations", available at https://www.unive.it/web/fileadmin/user_upload/dipartimenti/DMAN/pubblicazioni_scientifiche/working_papers/2025/2025wp02.pdf (last accessed 25 April 2026).

[179] OECD, "Mapping Relevant Data", 12.

[180] S. Kutschera, "Incidental Data: Observation of Privacy Compromising Data on Social Media Platforms" (2023) 4 International Cybersecurity Law Review 91.

[181] C. Fan et al., "A Review on Data Preprocessing Techniques toward Efficient and Reliable Knowledge Discovery from Building Operational Data" (2021) 9 Frontiers in Energy Research.

[182] Ibid.

[183] Krämer, Schnurr and Broughton Micova, "Role of Data", 110–11.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

at advancing scientific knowledge, informing policy decisions or providing public services.

However, the commercial/non-commercial distinction is far from clear-cut. A distinction may be drawn between *who* collects the data, referring to the institutional status of the entity as for-profit or not-for-profit, which itself is not entirely binary,[184] and *why* the data are collected, namely the motivation underlying the *activity*, which may be commercial or non-commercial. Accordingly, not-for-profit entities, such as research institutions,[185] may nonetheless engage in commercial activities, for example through technology transfer and licensing,[186] and data collected by government agencies may have secondary commercial uses.[187] However, as such revenues are ultimately intended to support overarching non-profit objectives, such as advancing scientific knowledge or public service,[188] these activities may be more accurately characterised as "semi-commercial", consistent with their hybrid organisational nature. As there is a range of hybrid entities, their activities likewise involve diverse degrees of semi-commerciality.[189]

In sum, the taxonomy of data encompasses various dimensions; however, the various categories are not dichotomous and often involve mixed contingencies. Understanding these aspects provides valuable insights regarding the underlying principles of data regulation.[190]

VI. A Matrix of Data-Sharing Regulation

Regulatory frameworks should rest on clear normative justifications,[191] achieved through the consistent application of principles and resistance to unfounded "logical" tendencies.[192] Accordingly, any recognition of proprietary rights in data must align with normative principles and data-sharing norms are no exception. To this end, we propose a matrix that may serve as a tool for critically examining existing regulations and may

[184] Mereghetti and Nuccio, "Towards a Taxonomy".
[185] Research institutions are diverse and their categorisation is complex: L. Cruz-Castro et al., "The Classification of Public Research Organizations: Taxonomical Explorations" (2020) 29 Research Evaluation 377.
[186] L. Jyh-An, "Licensing Open Government Data" (2016) 13 Hastings Business Law Journal 207; Hardy and Maurushat, "Opening Up".
[187] K. Debackere and R. Veugelers, "The Role of Academic Technology Transfer Organizations in Improving Industry Science Links" (2005) 34 Research Policy 321.
[188] Recommendation 2008/416/EC on the Management of Intellectual Property in Knowledge Transfer Activities (OJ 2008 L 146 p.19), Recital. 4; Annex I, art. 3.
[189] Mereghetti and Nuccio, "Towards a Taxonomy".
[190] For a similar view, see A.T. Kronman, "Mistake, Disclosure, Information, and the Law of Contracts" (1978) 7 Journal of Legal Studies 1, 13–14; Krämer, Schnurr and Broughton Micova, "Role of Data", 112.
[191] R. Slater, "The Regulation of Known Unknowns: Toward Good Regulatory Governance Principles" (2009) 3 Regulatory Governance Brief.
[192] J.M. Balkin, "Understanding Legal Understanding: The Legal Subject and the Problem of Legal Coherence" (1993) 103 Yale Law Journal 105, 114.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

also assist in building the backbone of future data sharing or access regulations. The proposed matrix does not prescribe specific type of regulation or concrete rules, but rather it functions as a roadmap, ensuring that emerging data governance models remain grounded in coherent theoretical foundations. Under the proposed matrix, the norm adopted should match the need to provide incentives for the collection and storage of data, which is the bedrock of any norm granting exclusive rights to data. In the absence of a need to incentivise these activities, the default should be a pro-access and non-proprietary approach to data. Regulatory frameworks pertaining to data apply to diverse situations, reflecting the complexity of the data ecosystem. Therefore, the various applicable frameworks under the proposed matrix comprise a range of norms, from a pure open-access and unrestricted data-sharing model at one end to a proprietary model that grants exclusive rights and control over data at the other. Under the open-access model, data holders would be obliged to allow access to data and its reuse, and data sharing may be based on a mandatory proactive basis. Between these extremes lie models which we term "calibrated access", addressing mid-range situations that are the most challenging. These include cases that may justify limiting data exclusivity through data-sharing obligations, at least to a certain degree, or softening mandatory data sharing by recognising limited exclusivity. Under the proposed matrix, the applicable models, ranging from fully proprietary, through calibrated access, to fully open access, should be determined according to the three questions – "what", "who" and "why" – used in the taxonomy above. Data taxonomy, therefore, aligns the adopted norms with key factual parameters.

As described above, the justification for applying a proprietary model to data lies in the need to provide an incentive for collection and storage of data. Therefore, under the proposed matrix, the most important factors are those associated with the question of "why" the data were collected, whether purposefully, as part of the collector's core activities, or incidentally, as a by-product and whether for commercial, semi-commercial or entirely non-commercial purposes. The commercial motivation to collect data may also indicate "who" the collector is – that is, whether the collector is a private or a public entity; however, the situations in this regard are complex and may be hybrid. As the intentions underlying data collection and the degree of commerciality of these activities vary across entities, the applicable regulatory model adjusts accordingly to reflect this complex range of scenarios. The types of data involved – that is, the "what" question – are significant in indicating the degree of personal interest in the data, which may affect its potential accessibility or portability. Additionally, the degree of data organisation may signal the intentions of the data holder.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

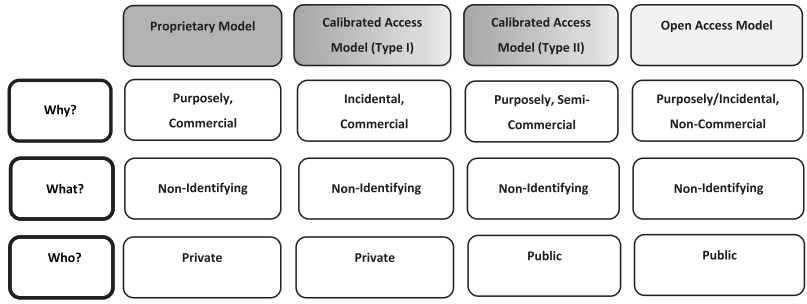


Figure 1. Proposed matrix of data regulation.

Figure 1 illustrates our proposed matrix of data-sharing regulation.

Under the proposed matrix, only data collected purposefully – that is, not incidentally as a by-product of another primary task and for commercial purposes – merit inclusion under the fully proprietary model because otherwise there may not be sufficient incentive for its collection and storage.[193] The incentive theory is primarily relevant to entities from the private business sector. An example is a commercial entity proposing services based on a dataset established and processed as part of its core business model.[194] The relevance of the incentive theory diminishes as we move away from these circumstances and the justification for applying a fully proprietary model that grants exclusive rights decreases.

At the second pole of the matrix, we propose adopting a fully open-access model as the normative default for non-commercial activities of data collection. This model is primarily relevant to public-sector entities. According to the proposed matrix, incidental collection of data, as a by-product of non-commercial activity, should be subject to a fully open-access model because there is no legitimate expectation of providing an incentive for data collection as a matter of policy under such a scenario. By the same principle, data purposefully collected in the course of public-sector non-commercial activities should be also subject to a fully open-access model because sharing does not undermine the motivation to collect the data. Datasets generated as part of the activities of public-sector entities belong to the public and should be accessible for the public good, subject to any other legal constraints such as privacy or national security obligations.[195] This captures the full-fledged meaning of the principle, according to which the default rule is a non-proprietary, open-access approach to data.

[193] Krämer, Schnurr and Broughton Micova, "Role of Data", 112.
[194] See e.g. LexisNexis, Dun and Bradstreet or Bright Data.
[195] For a similar discourse in the context of Freedom of Information, see I. Szekely, "Freedom of Information: A Constitutive Public Good in Democratic Societies" in A.S. Duff (ed.), *Research Handbook on Information Policy* (Cheltenham 2021).

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

Between the two poles lies a broad range of scenarios reflecting the complexity of data-collection activities, where the matrix "needle" shifts in response to the taxonomy's factors, thereby adjusting the applicability of the default non-proprietary, pro-access rule. For these mid-range situations, we therefore propose a calibrated access model that determines when it is justified to deviate from either pole. We identify two paradigmatic scenarios, arising in either private- or public-sector activities and accordingly term them "Type I" and "Type II". In general, Type I scenarios in the private sector support a move towards an open-access model, whereas Type II scenarios in the public sector support a move towards a proprietary model. However, the diversity and complexity of these situations require further fine-tuning within each type.

The Type I calibrated access model concerns the private sector and addresses incidental collection of data – that is, data collected as a by-product of another primary task that is carried out in the course of a commercial activity. Such data do not require applying a proprietary model because there is no need to provide an incentive for their collection and storage. A prominent example is the data collected by various online platforms providing services not directly associated with the data collected, such as social media services. While it is acknowledged that storing and managing such data involves non-trivial infrastructure costs, these are typically part of the broader operational model and not driven by the standalone value of the data itself at the time of collection.[196] As such, the normative justification for exclusive rights in such data may still be weaker than in cases where data collection is deliberate, targeted and value-driven. In these cases, the default rule of open access should prevail and mandatory data-sharing obligations should apply. In economic terms, granting exclusivity over this type of data may generate revenues beyond what is necessary to promote the activity (e.g. operating a social media service), thus reflecting a deadweight loss to the public good. Nevertheless, the commercial nature of the activity may justify some reservations in applying a full-fledged data-sharing obligation because the question of whether the data were collected and stored purposefully or incidentally may not always be easily answered. Therefore, we propose to adopt some exceptions to the general pro-access and data-sharing obligation rule for cases in which the data collected are part of the collector's core business model, which may reasonably be classified as data collected on purpose. The data products and services market is highly dynamic and

[196] Social media platforms primarily profit by collecting and analysing user data for advertising and targeting, as opposed to asserting exclusive rights over the publicly available data itself; see A. Gawer and N. Srnicek, "Online Platforms: Economic and Societal Effects", 10–15, available at https://www.europarl.europa.eu/RegData/etudes/STUD/2021/656336/EPRS_STU(2021)656336_EN.pdf (last accessed 27 April 2026).

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

the regulatory framework should be designed to accommodate complex and evolving business models, as it aims to promote data-driven innovation. Thus, the calibrated access model applicable to privately collected data, while adopting a pro-open-access rule, would nevertheless remain attentive to the broader commercial context in which these activities take place.

Further along the matrix, we propose adopting a second, the Type II calibrated access model, applicable to the public sector, which addresses cases in which the data are collected purposefully in the course of revenue-generating activities carried out by public-sector and not-for-profit entities, amounting to what may be characterised as semi-commercial activity. Although the not-for-profit status of the collecting entity may suggest the absence of economic incentives and thus support a stronger case for applying a fully open-access model, in practice many such entities purposefully collect data through activities that require significant investment of resources. In such cases, exclusivity over data may be justified as an exception to the general pro-access model, where it is necessary to provide an incentive for the activity. Universities offer a paradigmatic example. Although they operate primarily for non-profit purposes, they frequently engage in revenue-generating through technology transfer activities,[197] illustrating hybrid settings, combining societal missions with semi-commercial considerations.[198] In such cases, a deviation from a fully open-access model may be justified where control over the data is required to incentivise the activity.[199]

We recognise that not all public data is alike. Some public-sector entities engage in structured and resource-intensive data collection, such as national statistics, mapping or meteorological data and may operate under self-financing models. In such cases, full and immediate access may be balanced against the need to sustain the data infrastructure and differentiated approaches, such as access fees covering partial costs, may be appropriate. Where such entities pursue revenue maximisation rather than mere cost recovery, they may fall under the Type II calibrated access model, which captures public-sector entities operating in semi-commercial contexts, if incentives are required to sustain the activity. Thus, the calibrated access model applicable to publicly collected data moderates the pro–open-access rule in light of the hybrid, semi-commercial context in which these activities take place.

The proposed matrix is intended to adjust the general regulatory model – ranging from pro-proprietary to pro-open-access approaches – to the various

[197] Debackere and Veugelers, "Academic Technology Transfer".
[198] M. Holgersson and L. Aaboen, "A Literature Review of Intellectual Property Management in Technology Transfer Offices: From Appropriation to Utilization" (2019) 59 Technology in Society.
[199] M. Perkmann and K. Walsh, "University–Industry Relationships and Open Innovation: Towards a Research Agenda" (2007) 9 International Journal of Management Reviews 259.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

scenarios associated with the specific characteristics of data in each case. This matrix does not address the precise tailored norm applicable to the various scenarios, nor does it specify the concrete form or regulatory technique through which the rule should be crafted. Notably, the most challenging aspect of rule-making lies in designing pro-access and data-sharing obligations in cases where the data are in the possession of a data holder, particularly in the private sector, but also in hybrid public-sector scenarios. In such cases, regulation should not only establish that de facto possession does not create a legal right but also impose an affirmative obligation on the data holder to take active steps to enable data sharing. These scenarios may face conceptual resistance due to the common perception that de facto control over data implies a legal entitlement. However, once this conceptual barrier is overcome, the path is opened for developing regulatory obligations aimed at unlocking data. These regulatory techniques may evolve over time in response to legal and technological developments. The discourse on regulatory models for unlocking data is expanding, including debates over the pros and cons of coercive data-sharing rules (i.e. hard law) versus softer approaches such as "data spaces".[200]

VII. Advantages and Limitations of the Matrix

The proposed matrix offers a systematic and coherent basis for determining when exclusive rights over data are justified. It can be used as a tool to critically assess existing frameworks and may serve regulators and policy-makers as a roadmap for navigating the complex data environment while resisting stakeholder pressure to grant exclusivity. Property rights should not be acknowledged by inertia or, worse, by acceptance of a reality dictated by data possessors.

The proposed matrix is dynamic and technology-neutral, capable of adapting to future developments.[201] It considers a broad range of interests, including those of possessors, users and the public, facilitating access to information and fostering competitive markets. It thus provides a principled infrastructure for an efficient regulation of data.[202]

The proposed matrix is a skeletal model, designed to replace the prevailing assumption of vested exclusivity in data with a presumption of openness. Under this baseline, a proprietary legal model to data requires a robust justification. Shifting the default towards non-exclusivity could transform the data markets and promote the public good.

[200] European Commission, "Commission Staff Working Document on Common European Data Spaces", available at https://digital-strategy.ec.europa.eu/en/library/second-staff-working-document-data-spaces (last accessed 27 April 2026).

[201] S. Shadikhodjaev, "Technological Neutrality and Regulation of Digital Trade: How Far Can We Go?" (2021) 932 European Journal of International Law 1221.

[202] Drexl, "Designing", 260.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

Several objections may be raised. The key parameter, the "Why" question, asking whether the data were collected purposefully or incidentally as a by-product, is occasionally fuzzy and consequently may be subject to manipulations.[203] Data holders, particularly from the private sector, may claim that all data collected during their operations form part of their business model and therefore merit proprietary protection.[204] This reflects the proprietary impulse described earlier, whereby all commercially valuable outputs are perceived as owned assets. The answer lies in public policy, which supports depropertisation of data as the general rule. Physical control does not equate to ownership. Moreover, not all data collected by private entities are part of their core business.[205] For example, social media algorithms process only certain user-generated data, not all of it.[206] Claims by platforms over all data collected during their activities should therefore be rejected. While firms may legitimately monetise their activities, exclusivity over data should not rest solely on its inclusion within a business model. Evidentiary materials such as business documentation, purpose statements and actual usage patterns can help distinguish between purposefully and incidentally collected data.[207] Analogies may be drawn from other legal fields – such as tax, competition or contract law – where legal systems regularly differentiate between core and peripheral activities.[208]

Business models also evolve. Data collected without commercial purpose may later acquire value, yet proprietary rights should not be granted retroactively where no incentive for its collection was needed. Our argument is not that all data should be free or that private companies cannot profit from data, but rather that property rights must rest on principled justification, not mere technical control. Normative coherence in determining the boundaries of data ownership is essential and must be supported by clear regulatory or judicial mechanisms.[209]

A further objection is that the matrix may seem less progressive than the EU Data Act, which mandates data sharing for certain categories. Under the proposed matrix, data collected incidentally in the course of commercial activity would fall under the Type I calibrated access model, which deviates from the full-fledged proprietary model and mandates data

[203] Krämer, Schnurr and Broughton Micova, "Role of Data", 112.
[204] V. Grover et al., "Creating Strategic Business Value from Big Data Analytics: A Research Framework" (2018) 35 Journal of Management Information Systems 388.
[205] R. Van Loo, "Digital Market Perfection" (2019) 117 Michigan Law Review 815, 825.
[206] S. Ah Kim, "Social Media Algorithms: Why You See What You See" (2017) 2 Georgetown Law Technology Review 147.
[207] R.J. Allen and C.R. Callen, "The Juridical Management of Factual Uncertainty" (2003) 7 International Journal of Evidence & Proof 1.
[208] OECD, "Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations", 46, available at https://www.oecd.org/content/dam/oecd/en/publications/reports/2022/01/oecd-transfer-pricing-guidelines-for-multinational-enterprises-and-tax-administrations-2022_57104b3a/0e655865-en.pdf (last accessed 27 April 2026).
[209] Schweitzer et al., "Data Access".

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

sharing, while allowing limited exceptions in light of the commercial context. Critics may thus argue that the matrix steps back from the EU Data Act's more interventionist stance. However, the matrix is a principle-based model, designed to facilitate data sharing, while recognising justified cases for exclusivity. It offers a conceptual tool for assessing the legitimacy of any data law.

In this light, the matrix actually supports the EU Data Act's baseline rule that data generated and collected through the operation of IoT devices by commercial entities should be shareable. Yet, applying the matrix reveals that the EU Data Act's sharing obligation is limited – restricted mainly to data users rather than to other interested businesses.[210] Compliance with personal-data protection may partly explain this narrow scope. Similarly, the DMA requires gatekeepers to share data with originators (either business-to-business or business-to-customer), therefore it is not establishing a general pro-access regime in which data is open to all. The matrix helps expose the limited reach of these frameworks and encourages discussion of their normative basis.

The EU Data Act's trade-secret exception – the so-called "hand-brake mechanism", which allows data holders to withhold data that qualify as trade secrets[211] – also merits analysis under the matrix. Trade-secret protection applies to information that confers a competitive advantage derived from its secrecy. The EU Data Act mandates an objective, case-by-case framework for assessing trade-secret claims.[212] Is this exception supported by the matrix? The answer might be yes. Data that form part of the core business are typically collected purposefully, structured for added value and may qualify as trade secrets exempt from the sharing obligation. Data collected from connected vehicles to develop autonomous driving systems, is one example.[213] Conversely, data collected incidentally in the operation of IoT devices, often have no expected or feasible business advantage and therefore remain raw, without further organisation by its collector. Such data should be subject to a data-sharing model. Data from smoke detectors is an example.[214] While it could be argued that any data collected from connected devices and services may acquire commercial value simply by being inaccessible to others, this alone does not justify proprietary protection. Under the proposed matrix, the mere fact of secrecy is insufficient; exclusivity over data should only occur where there is intentional collection, purposeful structuring or demonstrable planned commercial use.

[210] Kerber, "Governance of IoT Data"; Schweitzer et al., "Data Access", 208–09, 215.
[211] EU Data Act, arts. 2–3.
[212] Ibid., art. 38.
[213] J. Drexl, "Data Access and Control in the Era of Connected Devices: Study on Behalf of the European Consumer Organisation", 41, available at https://www.beuc.eu/sites/default/files/publications/beuc-x-2018-121_data_access_and_control_in_the_area_of_connected_devices.pdf (last accessed 27 April 2026).
[214] Ibid.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press

Another major concern relates to enforcement feasibility. Critics may argue that requiring private entities to share data is impractical, echoing similar objections raised under the DMA's gatekeeper obligations.[215] Enforcement of data-sharing duties poses genuine challenges. However, enforcement difficulties have long accompanied intellectual property rights themselves, yet this has never precluded their recognition where justified. Similarly, data-sharing obligations can and should be accompanied by robust enforcement mechanisms, including regulatory oversight, penalties for non-compliance and clear procedural guidance.

VIII. Concluding Remarks

Contributing to ongoing discussions in the emerging field of data law, this article proposes a matrix that provides a principled framework for assessing when data exclusivity is warranted. It can be used to evaluate current policies, reveal inconsistencies and guide the design of future data governance. By shifting the default assumption from exclusivity to access, the matrix challenges the inertia of data propertisation and promotes a balanced approach, grounded in public interest and which seeks to avoid simplistic outcomes.

[215] Schweitzer and Metzger, "Data Access", 355.

https://doi.org/10.1017/S0008197326101585 Published online by Cambridge University Press