The Jean Monnet Center for International and Regional Economic Law & Justice

THE NYU INSTITUTES ON THE PARK

THE JEAN MONNET PROGRAM

J.H.H. Weiler, Director
Gráinne de Burca, Director

Jean Monnet Working Paper 1/22

Sergio Alonso De León

Intellectual Property law in the data economy: the problematic role of trade secrets and database rights for the emerging data access rights

NYU School of Law • New York, NY 10011
The Jean Monnet Working Paper Series can be found at
www.JeanMonnetProgram.org

ISSN 2161-0320 (online)
Copy Editor: Claudia Golden

New York University School of Law
New York, NY 10011
USA

Publications in the Series should be cited as:
AUTHOR, TITLE, JEAN MONNET WORKING PAPER NO./YEAR [URL]

Intellectual Property law in the data economy: the problematic role of trade secrets and database rights for the emerging data access rights

Sergio Alonso de León*

Abstract

The European Union (EU) has adopted a policy objective of fostering access to (non-personal) data. Law is one of the instruments towards this aim, particularly through the enactment of data access rights. The clash of intellectual property (IP) law with those rights is generally overlooked. This piece sheds light onto this crossroad.

In uncovering the interference of some IP rights with mandatory data access provisions, this article asks a series of questions in a much-needed conversation: what does data mean for the law? who 'has' the data, and who should get access to it? These questions have become increasingly consequential to our societies, and the law needs to tackle them comprehensively. Thus far, it has not. In this context, data-driven industries have quickly learned to operate in a landscape of legal uncertainty, articulating claims of exclusivity akin to property entitlements on data on the basis of specific IP instruments such as trade secrets and *sui generis* database rights. Whether these claims are legitimate is the puzzle this piece aims at disentangling.

* Senior Emile Noël Fellow, NYU School of Law, and member of the Legal Service at the European Parliament. I am grateful to Joseph H. H. Weiler, Gráinne De Búrca, Thomas Streinz, Katherine J. Strandburg and Jeanne Fromer. My insightful exchanges with them and the input received in the dynamic environment of the Jean Monnet Center have profoundly enriched my research. The views expressed are personal and do not represent any position of the European Parliament.

1. Introduction

Information has been valuable throughout human history. Today, the digital revolution has led to an unprecedented ability to capture and manage information in the form of data. In turn, data unlocks access to a new type of information (often in the form of metadata), whose value can yield vast benefits through increasingly sophisticated technologies. Let me illustrate this with an example. Online shopping might have transformed the retail industry, but perhaps more revolutionary still is the new information created in the process, such as the innumerable parameters of consumer behaviour in relation to a product, its pricing, the timing of the purchase or the associated purchases. Technology enables navigating these massive information points that did not exist or were unmanageable before. The challenges for societies on the amount of

information that becomes available with the data revolution are colossal, in particular when access to that information becomes increasingly unequal.[2]

I make two principal claims in this piece.

The first is that data is a new, different 'object of the law': neither a new resource nor, in itself, a creation of human ingenuity, which bears no 'private property' sign on it. The second is that data holders sometimes deploy arguments based on certain IP instruments to attain property-like entitlements on data, and that is problematic.

IP revolves around the idea of attributing a certain, limited ownership to the creator or innovator with a balancing of societal interests; that is, promoting social values and dissuading moral wrongs.[3] Two premises underpin the legitimacy of IP law instruments: a factual link between the beneficiary and the outcome of the creative or inventive process and a set of policy considerations. There is a misfit between these premises and the characteristics of data. I will explain why below: for now, suffice it to advance that, firstly, the attribution of data (to whom it should belong) is much more problematic than traditional IP instruments assume; and, secondly, the policy considerations are very different from those underpinning the different IP instruments. Classical IP law does not

[2] Some celebrated essays paint a frightening picture if the trend of unequal access to information brought about by data technologies continues. Harari (Yuval Noah Harari, *Homo Deus: A Brief History of Tomorrow* (Harper 2017)) reflects on the future 'dataism', where (artificial) intelligence grows and (human) consciousness diminishes with all organisms being just algorithms and life just data processing; Zuboff preconises an era of surveillance capitalism where human experience is reduced to free raw material (behavioural data) fed to AI that will anticipate what we will do and manipulate us at will (Soshanna Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (Profile Books 2019)); Hildebrandt posits that data-driven technologies undermine, reconfigure and overrule the ends of the law in a constitutional democracy, jeopardising law as an instrument of justice, legal certainty and the public good (Mireille Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* (Edward Elgar Publishing 2015)); Austin argues that data-driven technologies are 'world-making' since
they don't simply capture and represent a given world, but rather create the worlds they are designed to see (Lisa Austin, 'From Privacy to Social Legibility' (2022) 20 (3) Surveillance & Society 302); Pagano reveals the paradox that the knowledge-intensive characteristics of technology should favour a transparent democratic economy impregnated with non-rival knowledge; instead, big tech flourish on the successful seizure of information (Ugo Pagano, 'The Crisis of Intellectual Monopoly Capitalism' (2014) 38 Cambridge Journal of Economics 1409)).

[3] Jeremy Waldron, 'From Authors to Copiers: Individual Rights and Social Values in Intellectual Property' (1993) 68 Chicago-Kent Law Review 841.

provide the right framework for data and, although in most cases it is not applicable overall, some instruments within its contours interfere more severely than it is usually admitted.

This piece brings to fore the concept of data for the law and pinpoints its relevant characteristics, with the IP law and competition law framings in the background. It reflects on the lack of direct mechanisms for 'owning' the data, and consequently the workarounds of the data holders to retain control. Then, it lays down the context for the recent enactment of data access rights by the EU, aiming at the desired policy objective of fostering data sharing. Against this backdrop, it identifies the IP instruments susceptible to interfering with such data access rights, which are DB and TS rights. After a contextual analysis of each of them, I conclude that leveraging DB and TS rights to assert control over data oversteps the original justification for these instruments. In a time where the EU is in the process of enacting a framework for the data economy – aiming at more data in circulation and a fairer digital landscape for consumers and smaller operators – it is critical to elucidate this blind spot.

2. Framing the concept of 'data' and the goal of data sharing

A. Rethinking the notion of data

What is data? Why does it matter?

For our purposes, data is 'representation of information in digital form'.[4] Traditionally, one could separate within the notion of information:[5] a) the communication channel, b) the way it is recorded (syntactic level)[6] and c) the meaning or content (semantic level). Strictly speaking, data equates only to the syntactic level of information. The disambiguation between data and information is pertinent for us[7] because while the law can sometimes protect certain information, this is not always equivalent to a protection of the data,[8] something particularly worth underlining in the field of IP.

Data is under the spotlight for good reasons. It constitutes the fundamental building block of today's information society.[9] As machine learning has matured, artificial intelligence (AI) has exploded,[10] making data the critical factor for technological development.[11] Thanks to data, AI systems move beyond instantiations of human logic to an unimagined

[4] The new EU laws on data agree on the definition of data as 'any digital representation of acts, facts, information and any compilation of such acts, facts or information, including in the form of sound, visual or audiovisual recording', according to Parliament and Council Regulation (EU) 2022/868 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) [2002] OJ L152, art 2(1); Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act) [2002] OJ L265, art 2(24); and Commission, 'Proposal on a Data Act' COM (2022) 68 final, art 2(1). In earlier pieces, this relation to information was sometimes equivocal, for instance, Parliament and Council Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation (GDPR)) [2016] OJ L119, art 44 defines personal data as any information relating to an identified or identifiable natural person.

[5] Herbert Zech, 'Information as Property' (2015) 6 Journal of Intellectual Property, Information Technology and Electronic Commerce Law 192.

[6] In the digital context, this equates to machine-readable code which takes the form of 'electrical impulses in a binary code', see Georg Heath, 'Origins of the Binary Code' (1972) 227(2) Scientific American 76.

[7] 'Data' stems etymologically from the plural of the Latin *datum* or a 'given' fact; the original meaning 'pieces of information' shifted historically to become 'the outcome of a process'. See Daniel Rosenberg, 'Data Before the Fact' in L Geitelman (ed), *Raw Data is an Oxymoron* (MIT Press 2013) 14.

[8] See in particular below in the context of TS, section 3.C.

[9] Ivan Stepanov, 'Introducing a Property Right over Data in the EU: The Data Producer's Right – An Evaluation' (2019) 34(1) International Review of Law, Computers & Technology 65.

[10] Niklas Kühl, Marc Goutier, Robin Hirt, and Gerhard Satzger, 'Machine Learning in Artificial Intelligence: Towards a Common Understanding' in Tung Bui (ed), *Proceedings of the 52nd Hawaii International Conference on System Sciences* (ScholarSpace 2019) <http://hdl.handle.net/10125/59960> accessed 20 September 2022.

[11] Kai-Fu Lee, *AI Superpowers: China, Silicon Valley, and the New World Order* (Mariner Books 2018) 55, explains 'the invention of deep learning means that we are moving from the age of expertise to the age of data. Training successful deep-learning algorithms requires computing power, technical talent, and lots of data. But of those three, it is the volume of data that will be the most important going forward. That's because once technical talent reaches a certain threshold, it begins to show diminishing returns. Beyond that point, data makes all the difference. Algorithms tuned by an average engineer can outperform those built by the world's leading experts if the average engineer has access to far more data.'

territory,[12] where economic opportunities abound[13] and social disruptions occur. Data not only permits the management and monetisation of vast arrays of information,[14] whose cost is approaching nil,[15] it also increases usable information by extracting relevant inferences from metadata. The Internet of Things (IoT) with its ubiquitous sensors is capturing the physical world in an unprecedented way. The relevance of data (and metadata) regulation through law is thus immense.[16]

Data is called our era's 'new oil' so as to highlight its importance,[17] but the comparison is misleading and, arguably, far from innocent.[18] Information can be found in nature, but data results from its transformation into a digital representation and, therefore, cannot be extracted 'raw' from nature.[19] The metaphor 'data is just another resource' undermines the urgency for new rules to govern it. Such framing conveys the idea that existing rules can be adapted to the peculiarities of data, rather than data remaining in a limbo and calling for the legislator to intervene and enact specific rules for data in the sectors where it is necessary.

What are the characteristics of data?

[12] David Lehr and Paul Ohm, 'Playing with the Data: What Legal Scholars Should Learn About Machine Learning' (2017) 51 University of California Davis Law Review 653, 717.
[13] OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being* (OECD Publishing 2015), 20 <https://www.oecd.org/sti/data-driven-innovation-9789264229358-en.htm> accessed 20 September 2022.
[14] Russel Ackoff, 'From Data to Wisdom' in *Ackoff's Best* (John Wiley 1999) 170.
[15] Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think* (Houghton Mifflin Harcourt 2014) 29.
[16] Thomas Streinz, 'The Evolution of European Data Law' in Paul Craig and Gráinne de Búrca (eds), *The Evolution of EU Law* (3rd edn, OUP 2021) 907.
[17] Lauren Scholz, 'Big Data is Not Big Oil: The Role of Analogy in the Law of New Technologies' (2019) 86(4) Tennessee Law Review 863 describes how the analogy was used for the first time in 2006 and became commonplace in the debates on the matter, being picked up by regulators as apt to describe the new reality. Despite controversies, comparisons to oil continue to be put forward: see for example, Marcin Szczepański, 'Is Data the New Oil?' (2020) *European Parliamentary Research Service Briefing (PE 646.117)* <https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/646117/EPRS_BRI(2020)646117_EN.pdf> accessed 20 September 2022.
[18] Angelina Fisher and Thomas Streinz, 'Confronting Data Inequality' (2021) 1 IILJ Working Paper ('recourse to these metaphors depoliticises the process by which data is created').
[19] Danah Boyd and Kate Crawford, 'Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon', (2012) 15(5) Information, Communication & Society 662.

The starting premise is that data is conceptually different from other objects of the law.[20] This matters particularly in the context of IP law, as well as competition law. Linking with the previous considerations, the characteristics of data appear more clearly when contrasted with those of a natural resource; data it is not consumed when used and it is non-rivalrous.[21] This element is fundamental in order to understand the rationale of data sharing. The possible parallel use of data signifies that the cost of sharing is lower than for other assets, which adds to the low technical cost of transferring data. Consequently, the threshold for regulatory interventions in terms of mandatory access is lower and the claims of exclusivity harder to justify. This is particularly consequential in the field of IP law, because there is no tragedy of the commons,[22] and no crowding problem.

Other characteristics make data unique. Its general-purpose use,[23] because data collected with a specific objective can later serve other goals, which increases its potential societal benefits – a positive externality making a public intervention to stimulate sharing easier to justify.[24] The different means of supply, as the technical possibilities are countless (static copy, real-time access, with or without metadata, with or without possibility of subsequent transfer, data pooling, etc).[25] Its (at least potential) abundance, in that data is valuable not because it is scarce, but because it is ubiquitous and economies of scale make it profitable when there is a lot of it.[26]

---

[20] Susan Aaronson, 'Data Is Different, So Policymakers Should Pay Close Attention to Its Governance' in M Burri (ed), *Big Data and Global Trade Law* (CUP 2021) 340.

[21] Néstor Duch-Brown, Bertin Martens and Frank Mueller-Langer, 'The Economics of Ownership, Access and Trade in Digital Data' (2017) 1 JRC Digital Economy Working Paper 1, 12; Nils-Peter Schepp and Achim Wambach, 'On Big Data and Its Relevance for Market Power Assessment' (2016) 7(2) Journal of European Competition Law & Practice 120, 121. This characteristic should come with the disclaimer that the physical infrastructure through which data is generated and moved cannot be said to be non-rivalrous. Likewise, there are also legal barriers to the use such as the rules applicable to personal data; see Marc Bourreau, Alexandre De Streel and Inge Graef, 'Big Data and Competition Policy' (2017) CERRE Report 15.

[22] Katharina Pistor, 'The Code of Capital: How the Law Creates Wealth and Inequality' (Princeton University Press 2019) 109.

[23] This is why the principle of data minimisation in the GDPR bears such a burden for data sharing and why many data-sharing strategies exclude personal data from its scope.

[24] Dan Ciuriak and Maria Ptashkina, 'The State Also Rises: The Role of the State in the Age of Data' (2020) Ciuriak Consulting Conference Paper 1.

[25] American Law Institute and the European Law Institute, 'Principles for a Data Economy' (2021), Chapter B <https://www.principlesforadataeconomy.org/> accessed 20 September 2022 includes several modalities of contracts for supply or sharing of data.

[26] Katharina Pistor, 'Rule by Data: The End of Markets? (2020) 83 Law and Contemporary Problems 101, 106.

Finally, but not less importantly, a coat of legal uncertainty surrounds data. The mere discerning of the sources of governance proves a daunting task.[27] Law governs data only partially. It does so together with a mix of technical codes, standards, protocols and contracts of diverse kinds. This emerging new 'order' is exceptionally fragmented[28] and, while data is global in its essence and easily movable across borders, the regulation of data can only stem from national or regional jurisdictions, with no promising multilateral approach in sight.[29]

Despite these characteristics, and in this the analogy with natural resources is revealing, there is a rush to extract, to accumulate and to develop new methods to monetise data,[30] often to the exclusion of others.[31] The incentives for the exploitation of the competitive advantages inherent in the refusal to share are too evident.[32] Additionally, data being so vast, some data holders could be in a situation where they do not know what information they possess, and might fear it could prove harmful for them if shared with others. In this context, the patchy legal framework applicable to data enhances the temptation for data holders to leverage existing legal tools, as we shall see, such as TS and DB rights to attain exclusivity.[33]

What determines control over data?

[27] Thomas Streinz, 'International Economic Law's Regulation of Data as a Resource for the Artificial Intelligence Economy' in Shin-yi Peng, Ching-Fu Lin, T Streinz (eds), *Artificial Intelligence and International Economic Law: Disruption, Regulation, and Reconfiguration* (CUP 2021) 176.
[28] Douglas W Arner, Giuliano Castellano, Erik Selga, 'The Transnational Data Governance Problem' (2022) Berkeley Technology Law Journal (forthcoming).
[29] Aaronson (n 19) 349.
[30] Jathan Sadowski, 'When Data is Capital: Datafication, Accumulation, and Extraction' (2019) 6(1) Big Data & Society 1.
[31] Alberto Alemanno, 'Big Data for Good: Unlocking Privately-Held Data to the Benefit of the Many' (2018) 9 European Journal of Risk Regulation 183; and Sebastian Lohsse, Reiner Schulze and Dirk Staudenmayer, 'Introduction' in Lohsse, Schulze and Staudenmayer (eds), *Trading Data in the Digital Economy: Legal Concepts and Tools* (Nomos 2017), 15.
[32] Joseph Stiglitz, 'The Revolution of Information Economics: The Past and the Future' (2017) 23780 NBER Working Paper 1, 6.
[33] Dan Ciuriak, 'The Economics of Data: Implications for the Data-Driven Economy' in Rohinton P Medhora (ed), *Data Governance in the Digital Age* (Centre for International Governance Innovation 2018) 14.

Data being a new, distinct object of the law, there are three ways in which one could conceive law-supported mechanisms of control over data:

1) The first would be to establish *de lege ferenda* specific property rights on data. There has been an intense scholarly debate on the idea,[34] but a consensus is emerging that this would be inadequate for incentivising data sharing[35] and the proposals have failed to gain traction.[36]

2) The second would be to rely on the infrastructure through which data is captured, managed and transferred. This consists both of the physical infrastructure (protected by traditional property on goods), and the *software* (if not open source, eventually protected by copyright).

3) The third would be to assert the use of existing IP rights in order to substantiate claims on data and to legitimise exclusivity over it.[37] Without specific changes in the legal

[34] The concept of *Datennutzungsrecht* was constructed and debated notably among German academia. Wolfgang Kerber, 'A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis' (2016) 11 Gewerblicher Rechtsschutz und Urheberrecht Internationaler Teil 3; Michael Dorner, 'Big Data und "Dateneigentum" Grundfragen des modernen Daten-und Information-shandels' (2014) 30(9) Computer und Recht 617; Thomas Hoeren, 'Big Data and the Ownership in Data: Recent Developments in Europe' (2014) 36 European Intellectual Property Law Review 751; Herbert Zech,'Data as Tradeable Commodity' in Alberto de Franceschi (ed), *European Contract Law and the Digital Single Market* (Insentia 2016) 51; Jeffrey Ritter and Anna Mayer, 'Regulating Data as Property: A New Construct for Moving Forward' (2018) 16 Duke Law & Technology Review 220, 253; Daniel Zimmer, 'Property Rights Regarding Data' in Lohsse, Schulze and Staudenmayer (n 30) 106; Lothar Determann, 'No One Owns Data' (2018) 70 Hastings Law Journal 1, 5; Stepanov (n 8). Data ownership has also been subject to intense debate in China; however, for fear of stifling both innovation and competition, the idea of creating a property framework for data has also been unsuccessful: see Celine Yan Wang, 'Governing Data Markets in China: From Competition Litigation and Government Regulation to Legislative Ordering' (2022) 13(1) George Mason International Law Journal 1.

[35] Alek Tarkowski and Francesco Vogelezang, 'The Argument Against Property Rights in Data' (2021) Open Future Policy Brief 1, 9; Josef Drexl, 'Designing Competitive Markets for Industrial Data: Between Propertisation and Access' (2017) 8 Journal of Intellectual Property, Information Technology and Electronic Commerce Law 257, 291.

[36] Bernt Hugenholtz, 'Data Property in the System of Intellectual Property Law: Welcome Guest or Misfit?' in Lohsse, Schulze and Staudenmayer (n 30) 76. The European Commission seemed open to this idea in its 2015 Communication, 'A digital single market strategy for Europe', COM (2015) 192 final, stating that it would address 'the emerging issues of ownership, interoperability, usability and access to data'. The subsequent Communication of 2017, 'Building a European data economy', COM (2017) 9 final, no longer mentions the idea of ownership.

[37] Ana Nordberg, 'Trade Secrets, Big Data and Artificial Intelligence Innovation: A Legal Oxymoron?' in Jens Schovsbo, Timo Minssen and Thomas Riis (eds), *The Harmonization and Protection of Trade Secrets*

framework, there has been, in recent years, 'uncoordinated but self-interested efforts of the data operators and their lawyers to articulate legal mechanisms for the effective control of data',[38] which can be presented as a quiet revolution in the legal status of data towards a '(de facto if not de iure) proprietary informational property'.[39]

Against this backdrop emerges the essential interrogation of this paper: whether we want to underpin the existing second layer with a third layer of control over data; in other words, whether our legal order should endorse these 'legal walls' on data, should be a well-calibrated choice by society and reflected in purposive rules. As anticipated, I think that the considerations underpinning IP law are largely inapplicable to data. My claim is that operators with the technological power to extract data might well decide to materially exclude others from it, but should not do so on the basis of additional legal protections that were not intended for the data economy, and only inasmuch as the legislator does not decide otherwise in specific circumstances by granting data access rights and data sharing obligations. The relevance of this question emerges as the new data access rights would certainly prevail over the second layer of control, but would have a more ambiguous relationship with the third layer.

B. The elusive objective of data sharing. Policy considerations: protect personal data and foster data sharing

Why should we promote (or not) data sharing?

The positive effects of increasing the volume and quality of data are well documented, from an economic[40] and technological perspective.[41] The estimation of the added value of

in the EU: An Appraisal of the EU Directive (Edward Elgar 2020), 197; Amy Kapczynski, 'The Law of Informational Capitalism' (book review) (2020) 129 Yale Law Journal, 1460, 1499.

[38] Julie Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism* (OUP 2019) 20.

[39] Julie Cohen, 'Law for the Platform Economy' (2017) 51 University of California, Davis 133, 154.

[40] OECD (n 12), 131 provides evidence on how data drives economic value creation and fosters new products, services and industries.

[41] For the effects on technological development, see Ajay Agrawal, Joshua Gans and Avi Goldfarb, 'Economic Policy for Artificial Intelligence' (2019) 19 Innovation Policy and the Economy 139. With an economic focus, see the often quoted article by James Manyika, Susan Lund, Jacques Bughin, Jonathan Woetzel, Kalin

data to the economy is stratospheric[42] and the benefits of data collection are compounded when data access, data sharing and data re-use are widespread,[43] which in turn promote further innovation and a fairer playing field for smaller operators.[44]

Against this backdrop, two principal concerns arise.

The first concern is that not all data should be shared. This relates fundamentally to personal data – a vast and dynamic sphere that I need to keep outside the focus of this research, but which reflects an increasing preoccupation with the undermining of our societal[45] and democratic standards.[46]

The second concern is that access to non-personal data is increasingly concentrated,[47] raising worrying prospects for the fairness of the economy.[48] Evidence is mounting that the economic incentives nudging data operators to avoid sharing data exceed other incentives. As competition authorities have detected a few years ago, data exclusivity is turning into an instrument of market power, with significant risk of abusive conduct

---

Stamenov and Dhruv Dhingra, 'Digital Globalization: The New Era of Global Flows' (McKinsey Global Institute March 2016) available at <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/digital-globalization-the-new-era-of-global-flows> accessed 20 September 2022.

[42] 'The European Data Market Monitoring Tool: Final Study Report' by the European Commission, the Lisbon Council and International Data Corporation, available at <https://datalandscape.eu/sites/default/files/report/D2.9_EDM_Final_study_report_16.06.2020_IDC_pdf.pdf> accessed 20 September 2022, estimates that the value of the data economy, which measures the overall impacts of the data market on the economy as a whole, exceeded the threshold of 400 billion Euro in 2019 for the EU27+UK and is expected to grow at a higher pace in the coming years.

[43] The OECD estimates that data access and sharing can help generate social and economic benefits worth between 0.1 and 1.5 per cent of GDP in the case of public-sector data, and between 1 and 2.5 per cent of GDP when also including private-sector data. OECD, *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use* (OECD Publishing 2019) <https://www.oecd.org/digital/enhancing-access-to-and-sharing-of-data-276aaca8-en.htm> accessed 20 September 2022..

[44] Jacques Crémer et al, 'Fairness and Contestability in the Digital Markets Act' (2019) Yale Tobin Center for Economic Policy Discussion Paper No 3 14, 16.

[45] Among many other dedicated studies to the concern regarding the use of personal data by digital corporations, Zuboff denounces the collection and repurposing of data for means social control, in the absence of citizens' awareness or means of combat. Zuboff (n 1) 53.

[46] Gerard de Vries, 'Do digital technologies put democracy in jeopardy?'. in Mireille Hildebrandt and Kieron O'Hara (eds), *Life and the Law in the Era of Data-Driven Agency* (Edward Edgar, 2020) 135.

[47] Fisher and Streinz (n 17).

[48] Pistor (n 21) 204 uncovers a trend to enclose the commons and capture its monetary rewards. Rory Van Loo, 'The Missing Regulatory State: Monitoring Businesses in an Age of Surveillance' (2019) 72 Vanderbilt Law Review 1563 underlines the irony of the information age where companies responsible for extensive surveillance of individuals remain opaque.[48]

through the exploitation of data collection capabilities to foreclose the market,[49] a scenario exacerbated by the rapid evolution of digital markets and the multi-purpose nature of the power of data operators.[50]

The policy objective of data sharing in the EU

The policy objectives and concerns discussed above are common to all countries with technological ambitions, but several aspects characterise the EU as a particularly relevant jurisdiction. Firstly, the EU has made high-privacy standards its flagship in digital policy.[51] Secondly, it has prioritised explicitly a vision of fostering data sharing, the so-called European Strategy on Data (ESD).[52] The third is that, perhaps aware of its weaknesses in terms of locally brewed technological development,[53] it has doubled down as a regulatory powerhouse[54] and is pioneering data laws that other technologically advanced jurisdictions are yet to consider.[55]

---

[49] Autorité de la concurrence and Bundeskartellamt, 'Competition Law and Data' (2016) <https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.pdf?__blob=publicationFile&v=2> accessed 20 September 2022.

[50] Nicolas Petit, *Big Tech and the Digital Economy: The Moligopoly Scenario* (OUP 2020).

[51] Although this article brackets off privacy rules, it is pertinent to flag that there seems to be an increasing appetite to have a reasonable discussion as to the extent in which personal data protection can be both efficient and not a burden to data sharing. Examples of this statement can be found in the enhancement of the portability right of the GDPR by the recent proposal for a Data Act (n 4), and a less categorical case-law, more accepting of a risk-based approach to anonymised data (Case C-582/14 *Breyer v Deutschland* (ECLI:EU:C:2016:779)), and an increasing concern about the GDPR becoming the 'law of everything' (opinion of Advocate General Bobekin, Case C-245/20 *X and Z v Autoriteit Persoonsgegevens* (ECLI:EU:C:2021:822)].

[52] Commission Communication, 'European Strategy on Data' COM (2020) 66 final, whose main goal is improving both the quantity and quality of data in circulation, its accessibility and use. It follows up on previous documents, notably in 2017 'Building a European Data Economy' (n 35); in 2015 'A Digital Single Market Strategy' (n 35); and in 2014 Commission Communication, 'Towards a Thriving Data-Driven Economy' COM (2014) 442 final. The objective is fully supported by the other institutions: see European Council conclusions of 2 October 2020 (EUCO 13/20) and, for European Parliament, ITRE report on a European strategy for data of 3 March 2021 (2020/2217(INI).

[53] 'European Strategy on Data' (n 51) 3.

[54] Anu Bradford, *The Brussels Effect : How the European Union Rules the World* (OUP 2020), xiv.

[55] Some observers have stressed that enacting rules on data is indeed the best approach to steer the development of AI technology. See Alicia Solow-Niederman, 'Administering Artificial Intelligence' (2020) 93 Southern California Law Review 633.

Respecting this particular equation of increasing data access through regulation while keeping an elevated level of personal data protection,[56] the EU policy makers are convinced that an uptake in data sharing will boost competitiveness, fairness and innovation – with synergies in key sectors such as environmental protection or health care –[57] and will be an essential factor for regaining European digital sovereignty,[58] thus healing the ills described above.

Yet the objective remains elusive.[59] There are three sets of reasons why data generated is not shared. The first is that there are regulatory restrictions on the sharing of data, which relates primarily to the personal data protection rules. The second relates to market inefficiencies such as data collectors not finding data markets to place their data or the data not being interoperable. The third is that data operators decide to hold on to their data.

As indicated, considerations as to the first set of reasons are beyond the intended scope of this piece. As to the second set of reasons, the EU has tried to encourage the growth of efficient and interoperable data markets through several recent pieces of legislation. They include the Regulation on the Free Flow of non-personal Data (facilitating self-regulation

[56] This scheme diverges fundamentally from the paradigm of free data flows referred to as 'Silicon Valley consensus', because one of the tenets of the EU set of policy prescriptions is stringent privacy regulation. See Thomas Streinz, 'Digital Megaregulation Uncontested? TPP's Model for the Global Digital Economy' in B Kingsbury et al (eds), *Megaregulation Contested: Global Economic Ordering After TPP* (OUP 2019) 337.
[57] See for example, Massimo Russo, David Young, Tian Feng and Marine Gerard, 'Sharing Data to Address Our Biggest Societal Challenges' (2021) BCG Henderson Institute <https://www.bcg.com/publications/2021/data-sharing-will-be-vital-to-societal-changes> accessed 20 September 2022 and ALLEA, EASAC and FEAM, 'International Sharing of Personal Health Data for Research' (2021) European Academies Science Advisory Council < https://easac.eu/publications/details/international-sharing-of-personal-health-data-for-research/ > accessed 20 September 2022.
[58] Digital sovereignty refers to the EU's ability to act independently in the digital world both with protective mechanisms and offensive tools to foster digital innovation. See European Commission, European Political Strategy Centre, 'Rethinking Strategic Autonomy in the Digital Age' (EU Publications Office 2019). The Report on Artificial Intelligence in a Digital Age (AIDA/9/04886) (2020/2266(INI)), point 158, 'finds that a lack of legal certainty, access to and sharing of high-quality data ... have led to a situation in which the EU competitiveness is constantly decreasing' relative to the AI-frontrunners China and the US.
[59] In her 2020 state of the Union address, Commission President Vor der Leyen underscored that about 80 per cent of the collected industrial data is never used: see <https://ec.europa.eu/commission/presscorner/detail/ov/SPEECH_20_1655> accessed 20 September 2022.

and enhancing cybersecurity),[60] the Open Data Directive (encouraging the development of open-source software and increasing the availability of public sector data),[61] the Data Governance Act (fostering trust in data intermediaries),[62] and the recent proposal for a Data Act (detailing rights and obligations of data users).[63] The rationale for these instruments is that a clearer framework and certain incentives in the law will lead to a growing exchange of data among operators. The reflections, for now abandoned, on the creation of a special property right on data examined before also bear upon this set of reasons.[64]

The emphasis of the legislator has been notoriously smaller as regards the third set of reasons, but there is a shift toward regulatory action in this regard.[65] A mapping of the interests at stake reveals that data holders often have the incentive to maintain exclusivity over their data, clinging to legal instruments to fortify their competitive positions,[66] in opposition to willing data re-users (particularly start-ups and Small and Medium Enterprises) who would benefit from an opportunity to enter and thrive in markets previously closed to them or dominated by a few big players.

There is increasing realisation that this is also an important factor explaining why data is not shared. The recent Digital Services Act (DSA)[67] and Digital Markets Act (DMA),[68] signal the regulatory momentum ignited by the perception that enclosure of data by big

---

[60] Parliament and Council Regulation (EU) 2018/1807 on a framework for the free flow of non-personal data in the European Union [2018] OJ 2018 L303/59.
[61] Parliament and Council Directive (EU) 2019/1024 on open data and the re-use of public sector information [2019] OJ 2019 L172/56.
[62] Data Governance Act (n 4).
[63] Draft Data Act (n 4).
[64] See section 2.A.
[65] Josef Drexl, *Data Access and Control in the Era of Connected Devices: Study on Behalf of the European Consumer Organisation BEUC* (BEUC (2018) 86 <https://www.ip.mpg.de/fileadmin/ipmpg/content/aktuelles/aus_der_forschung/beuc-x-2018-121_data_access_and_control_in_the_area_of_connected_devices.pdf> accessed 20 September 2022; Jason Furman et al, 'Unlocking Digital Competition: Report of the Digital Competition Expert Panel' (2019) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf> accessed 20 September 2022.
[66] Alemanno (n 30) 187.
[67] Regulation (EU) 2022/2065 of the European Parliament and the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act) [2022] OJ 277.
[68] Digital Markets Act (n 4).

tech is socially costly and that time is ripe for enacting a set of data access rights. Leaving aside privacy considerations, creating data access rights poses diverse challenges, one being interoperability, which the law could help to remediate,[69] and another being legal entitlements based on IP rights. We focus on the latter.

The blind spot: the interference of IP in the objective of data sharing

In the context just described, it is surprising that the digital agenda has remained agnostic as to the impact of IP laws in the goal of data sharing. The European Commission argued in 2017 that 'raw machine-generated data are not protected by existing IP rights', with negligible interference of the *sui generis* DB protection,[70] and no role for TS (under the assumption that it would lack commercial value).[71] The 2020 ESD mentions 'evaluating the IPR framework with a view to further enhance data access and use', mentioning both the DB and TS directives,[72] though the draft Data Act only touches upon the *sui generis* DB rights.[73] Besides this, none of the new EU laws on data affects IP rights.[74] Quite the contrary, they are deferential to existing IP rights. For example, the Data Governance Act clarifies that re-use of certain categories of data is conditional on the respect of IP and TS rights,[75] in a similar way as the draft Data Act.[76] An equivalent reproach can be made to the comprehensive 'principles for the data economy' put forward by a transnational

[69] Oscar Borgogno and Giuseppe Colangelo, in 'Data Sharing and Interoperability: Fostering Innovation and Competition through APIs' (2019) 35(5) Computer Law & Security Review 1, 3, explain how the creation of data access rights in the banking sector has contributed to the development of interoperable data infrastructures through APIs. Moreover, Jörg Hoffmann and Begoña Gonzalez Otero, in 'Demystifying the Role of Data Interoperability in the Access and Sharing Debate' (2020) Max Planck Institute for Innovation & Competition Research Paper No. 20-16, argue that the very existence of IPR on data has a bearing on the lack of interoperability.
[70] Building a European Data Economy (n 35) point 3.3.
[71] Commission, 'Commission staff working document on the free flow of data and emerging issues of the European data economy' [2017] SWD (2017) 2 final, 20.
[72] European Strategy on Data (n 51), 13.
[73] Draft Data Act (n 4) art 35. See below section 3.D..
[74] With the exception of the text and data mining exception incorporated in Parliament and Council (EU), Directive 2019/790 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC [2019] OJ L130, arts 3 and 4.
[75] Data Governance Act (n 4) 7 of explanatory memorandum, recitals 12 to 17 and art 5.
[76] Draft Data Act (n 4) art 8(6).

network of academic experts, which include a disclaimer in the rule on contracts for the transfer of data, indicating that IP rights, if applicable, would prevail.[77]

Why this deference? One possible explanation is the belief that there is no actual interference. Another would be a decision to maintain IP and TS protection despite the interference, an option of political economy that deliberately, and in my view mistakenly, avoids the problem. In either case, my understanding is that the consequences of certain IP instruments in the data world have not been sufficiently weighted in. I will argue through two case-studies, DB and TS protection, that this interference exists and that it is not justified.

3. Trade secrets and database rights: uninvited guests to the data economy

A. IP law in the data economy

The goal of incentivising innovation and creativity pervades IP laws. The essence of IP consists of letting creators and innovators appropriate the moral and economic benefit of their work, avoiding undesirable free-riders and other externalities (generally referred as 'tragedy of the commons').[78] IP is accepted under limitations such as a definite period after which protected objects become common goods. The technological context is relevant for the specific balance of interests, conditions and limitations embedded in the IP laws.[79] Indeed, under certain circumstances, IP instruments can stifle innovation, for instance, limiting follow-on innovation;[80] or be detrimental to other socially desirable

[77] Principles for a data economy (n 24), Principle 7. During the presentation of the principles on 19 October 2021, it was acknowledged that the IP framework is not prepared for the application of these or similar principles. See <https://principlesforadataeconomy.org/news-and-events/principles-for-a-data-economy-conference-2021/> accessed 20 September 2022.
[78] See Garrett Hardin, 'The Tragedy of the Commons' (1968) 162 Science 1243, 1246.
[79] The debate is not new. Technology has long brought about controversies as to whether new elements such as software merit classic IP protections such as, in this case, copyright. See for example, Richard H Stern, 'Symposium: The Future of Software Protection: The Bundle of Rights Suited to New Technology' (1986) 47 University of Pittsburg Law Review 1229.
[80] Yafit Lev-Aretz and Katherine J Strandburg, 'Privacy Regulation and Innovation Policy' (2020) 22 Yale Journal of Law & Technology 256, 307.

objectives.[81] The data economy brings many different factors into the picture, necessitating a reassessment of the core tenets of IP.

The traditional instruments of IP have built-in limitations that make them largely incapable of affecting data sharing. Copyright, even though omnipresent in the digital world, is not effective in enclosing data, because the copyright holder does not own the very data in which the copyrighted work is encoded.[82] The classical distinction of ideas/expression of ideas[83] mirrors the dichotomy of data/information explained before. Ideas become, after voluntary communication to others, 'free as the air to common use'[84] (equivalent to the semantic level of information), while the expression of them (equivalent to the syntactic level of information, or code) is copyrightable.[85] Patents do not raise concerns for the flow of data either, although they can become an indirect source of data enclosure as a derivative product of protection.[86]

Two instruments in the orbit of IP, BD and TS raise important challenges to the data economy. They are peculiar: DB protection because it manifestly does not concern

[81] Yafit Lev-Aretz and Katherine J Strandburg, 'Regulation and Innovation: Approaching Market Failure from Both Sides' (2020) 2 Yale Journal on Regulation Online Bulletin 2.
[82] Case C-128/11 *UsedSoft GmbH v Oracle International Corp* (ECLI:EU:C:2012:407) 42. Thomas Margoni and Martin Kretschmer, 'A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology' (2021) 7 CREATEe Working Paper Series, point out that the fact that a text and data mining exception was incorporated to the copyright directive proves that the potential interference, although small, can still arise. Drexl (n 64) points out a hypothetical interference by copyright protection for so-called application programming interfaces.
[83] See art 9(2) of the TRIPS agreement (Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, 1869 UNTS 299, 33 ILM 1197 (1994)) and art 2 of the 1996 World Intellectual Property Organization Copyright Treaty, which reads: 'Copyright protection extends to expressions and not to ideas, procedures, methods of operation or mathematical concepts as such.'
[84] Quote attributed to USSC Justice Brandeis. The often-cited US Supreme Court ruling *Harper & Row Publishers, Inc. v Nation Enterprises* [1985] 471 US 539, argued that 'copyright's idea/expression dichotomy strike[s] a definitional balance between the First Amendment and the Copy-right Act by permitting free communication of facts while still protecting an author's expression.'
[85] Zech (n 4) 194. This elementary distinction permeates EU legislation on informatics, for instance, Parliament and Council (EU), 'Directive 2009/24/EC on the legal protection of computer programs (Codified version) computer programs' [2009] OJ 2009 L111, art 1(2) delimits its scope to 'the expression in any form of a computer program' but not to 'ideas and principles which underlie any element of a computer program, including those which underlie its interfaces'. However, some authors have flagged that this distinction is not a complete guarantee of non-interference, as practice can reveal unintended creation of de facto exclusive rights on information per se (at the semantic level) while nominally protecting only the syntactic level of data. Tanya Aplin, 'Trading Data in the Digital Economy: Trade Secrets Perspective', in Lohsse, Schulze, Staudenmayer (n 30), 68.
[86] Drexl (n 64), 87.

creativity, and TS because its legal nature goes beyond many of the traditional parameters of IP.

B. The case for the database protection

The unique regime of database protection in the EU

Non-creative databases do not enjoy generalised international protection.[87] In the 1990s, when databases were becoming both increasingly sophisticated and easier to copy, a dilemma surfaced: to leave databases outside the scope of copyright, as the United States (US) did through the seminal ruling *Feist Publications*,[88] or to create a specific IP right similar to copyright, as the EU did through the DB Directive.[89] The EU intended to create a welcoming environment for the information market, but the idea has failed to attract significant interest from the industry.[90]

Defined as 'a collection of ... data arranged in a systematic or methodical way and individually accessible by electronic or other means', [91] databases can enjoy two types of

[87] Berne Convention for the Protection of Literary and Artistic Works (9 September 1886) 828 UNTS 221, art 2(5). At the level of the World Intellectual Property Organisation, the possible introduction of international protection of non-original databases has been discussed on several occasions, but the debates have never come to fruition.

[88] *Feist Publications, Inc v Rural Telephone Service Co,* (1991) 111 S Ct 1282. There were proposals to incorporate database rights within statutory law, but they never materialised. See Michael Freno, 'Database Protection: Resolving the US Database Dilemma with an Eye Toward International Protection' (2001) 34(1) Cornell International Law Journal 165, 167. This does not mean there are no other mechanisms to provide at least a partial protection, such as misappropriation rules: see Marshall Leaffer, 'Database Protection in the United States Is Alive and Well: Comments on Davison' (2007) 57 Case Western Reserve Law Review 855.

[89] Parliament and Council (EC) Directive 96/9/on the legal protection of databases [1996] OJ 1996 L77/20. On how this option remains relatively unique by global standards, see Jerome Reichman, 'Database Protection in a Global Economy' (2002) 16(2) Revue internationale de droit économique 455, 463.

[90] The Commission has carried out two evaluations of the database directive. The 2005 evaluation (DG Internal Market and Services Working Paper, First Evaluation of Directive 96/9/EC on the legal protection of databases < https://ec.europa.eu/info/sites/default/files/evaluation_report_legal_protection_databases_december_2005_en.pdf > accessed 20 September 2022, concludes that the sui generis right has failed to boost the creation of databases in Europe, and the 2018 Evaluation ((SWD (2018) 147 final) <https://ec.europa.eu/transparency/documents-register/detail?ref=SWD(2018)147&lang=en> accessed 20 September 2022, indicates that the economic benefits are unclear.

[91] DB Directive (n 88) art 1(2).

protection under EU law. One for creative databases akin to copyright protection,[92] and a second one, called *sui generis*, with no creativity requirement but with the added qualifying element of a 'substantial investment in either the obtaining, verification or presentation of the contents'.[93]

Only the second type is relevant for our purposes. The case law has delineated its contours as follows:

– In *Fixture Marketing*,[94] the Court of Justice of the European Union (CJEU) clarified the scope of DB protection, which includes the collection of data but not the individual data contained in it. Extraction and re-utilisation of individual data contained in a DB is, in itself, not precluded, except if it is a substantial part of the whole content of the DB.[95]

– In *British Horseracing Board*,[96] the CJEU defined 'substantial investment', which has 'to refer to the resources used to seek out existing independent materials and collect them in the database, and not to the resources used for the creation as such of independent material'.[97] There was no DB infringement where an individual set up a profitable website scrapping the data by the Horseracing Board (an operator devoting important resources to maintaining a DB on the performance of racing horses) as it was simply inputting the independently collected data.

– In *CV-Online Latvia*,[98] the CJEU added elements to the substantial investment threshold; namely, the re-use of data must preclude the DB holder from redeeming that

[92] DB Directive (n 88) art 3(1). The requirement of originality in a DB has been subject to a rather strict interpretation: see Case C-604/10 *Football Dataco v Yahoo! UK* (ECLI:EU:C:2012:115) 38. This type of DB, like traditional copyright, is does not raise particular concerns of interference with the data economy.
[93] DB Directive (n 88) art 7(1)..
[94] Case C-444/02 *Fixtures Marketing Ltd* (ECLI:EU:C:2004:697) 35, later confirmed in Case C-490/14 *Freistaat Bayern v Verlag Esterbauer GmbH* (ECLI:EU:C:2015:735) 20. On how this case law permits drawing the line between databases and 'data pools' see Andrea Ottolia, *Big data e innovazione computazionale* (Giappichelli Editore 2017) 76.
[95] DB Directive (n 88) art. 7(1) and Zech (n 33) 71.
[96] Case C-203/02 *British Horseracing Board Ltd and Others* (ECLI:EU:C:2004:128).
[97] ibid 31.
[98] Case C-762/19 *CV-Online Latvia v Melons* (ECLI:EU:C:2021:434). See further Sabine Jacques, 'CV-Online Latvia v Melons: In Search of Flexibilities Under the Database Directive' (*EU Law Live*, 8 June 2021)

investment.[99] It concluded that there can be an infringement when a search engine displays of third party's database and redirects users to the original website.

The risk of interference of the database protection with the data economy

Some observers conclude that this restrictive case law implies that the risk of interference with the data economy is low, as the collection and management of data by big data technologies would not pass the high 'substantial investment' threshold.[100]

However, the risk cannot be ruled out. Some experts have observed that the risk of interference has become more relevant with technological advancement.[101] The judicial interpretation separating the investment in collecting the data (not accounted as substantial investment) and the arrangement of the database (accounted as investment) lacks clarity in practice[102] and can be easily circumvented, for example via procedural systems to distinguish one investment and the other, or with technological protection measures to limit access to the data.[103] Actually, recent cases such as *CV-Online Latvia* law proves that, although the bar remains high, activities of meta-search websites are potentially infringing the *sui generis* right under certain conditions.[104]

<https://eulawlive.com/analysis-cv-online-latvia-v-melons-in-search-of-flexibilities-under-the-database-directive-by-sabine-jacques/#> accessed 20 September 2022.

[99] *CV-Online Latvia* (n 97) 46.

[100] Iain Connor, 'Database Rights are no "Impediment" to Europe's Data-Driven Economy' (*The Register*, 14 January 2016) <https://www.theregister.com/2016/01/14/database_rights_are_no_impediment_to_the_growth_of_europes_datadriven_economy_expert_says/> accessed 20 September 2022; Drexl (n 34) 268. Also in this line, Konrad Żdanowiecki, 'Recht an Daten' in Peter Bräutigam and Thomas Klindt (eds), *Digitalisierte Wirtschaft/Industrie 4.0* (Noerr LLP for BDI 2015) 21.

[101] Matthias Leistner, 'Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform' in Lohsse, Schulze, Staudenmayer (n 30) 27.

[102] Also, regarding the lack of clarity as to the epistemological distinction between creating and obtaining data in the case law, see Mark J Davison and P Bernt Hugenholtz, 'Football Fixtures, Horseraces and Spin-Offs: The ECJ Domesticates the Database Right' (2005) 27(3) European Intellectual Property Review 113.

[103] Jens Gaster, '"Obtinere" of Data in the Eyes of the ECJ: How to Interpret the Database Directive After British Horseracing Board Ltd et al v William Hill Organisation Ltd' (2005) 129 Computer und Recht: International 135.

[104] See Case C-202/12 *Innoweb BV* (ECLI:EU:C:2013:850) 40–41; Case C- 304/07 *Directmedia Publishing GmbH* (ECLI:EU:C:2008:552) 33 and *CV-Online Latvia* (n 97) 37.

Case law at national level confirms these suspicions of interference. In the German *Autobahnmaut* case,[105] a highway toll collector had set up a dynamic DB for billing users. An internet service provider involved in the toll payment system and with access to the data made it available to the transport companies who wanted to monitor their costs. The German Federal Supreme Court found the data sharing in breach of the *sui generis* DB right, therefore granting the toll collector a monopoly on the data generated by the users. This interpretation, which the German judges considered in conformity with the case law of the CJEU, implies that many situations in the context of machine-generated data would be covered by the *sui generis* DB right.[106]

The Commission has been hesitant when assessing the risk of interference, and its reading of the situation has fluctuated. In its 2005 evaluation of the DB Directive, it expressed concerns that the definition did not provide sufficient legal certainty, and it was 'precariously close to protecting data as property'.[107] The subsequent restrictive case law helped to appease the fears of interference with the data economy and in its 2018 evaluation[108] it concluded that the *sui generis* right does not apply broadly to the data economy. This consideration was underpinned by so-called 'spin-off theory' (databases that are by-products of another main activity of the maker should not enjoy protection). Adding this factor is questionable, first because it has no support in the case law, and second because it does not exclude a number of automatically generated databases from finding protection.[109]

The reflection announced by the 2020 ESD on the impact of the DB Directive concluded with a half-hearted conclusion in the 2022 draft Data Act. Article 35 of the draft Data

[105] Federal Supreme Court (Bundesgerichtshof) of 25 March 2010, Case I ZR 47/08 *Autobahnmaut* [2010] Gewerblicher Rechtsschutz und Urheberrecht 1004.
[106] Drexl (n 64) 106.
[107] 2005 Evaluation (n 89) 24.
[108] 2018 Evaluation (n 89) 12. One year before, it announced that it would evaluate the impact of the sui generis right on machine-generated data in its 2017 'Building a European Data Economy' (n 35) 10.
[109] Paolo Burdese, 'AI-Generated Databases: Do the Creation/Obtaining Dichotomy and the Substantial Investment Requirement Exclude the Sui Generis Right Provided for under the EU Database Directive? Reflections and Proposals' (17 December 2020) WIPO Academy, University of Turin and ITC-ILO – Master of Laws in IP – Research Papers Collection – 2019–2010, 10.

Act[110] includes a 'clarification'[111] that the 'sui generis right does not apply to databases containing data obtained from or generated by the use of a product or a related service'. This provision reveals a concern about the issue but the solution is not fully satisfactory. Firstly, because it assumes that DB rights never protected automatically (or IoT-) generated data, thus only 'clarifying' the matter. As explained above, if the DB meets the 'substantial investment' threshold, it fell within the scope of the directive, so the assumption is faulty. Secondly, because a 'clarification' of a directive by a regulation does not call for a modification of the national implementing laws, leaving a vacuum (EU law would not regulate such type of DB but Member States could), which could generate additional legal uncertainty.[112]

In conclusion, while DB protection has not emerged as an obvious obstacle to the flow of data in the EU – thanks to the little interest by the industry as well as the restrictive case law – it brings about uncertainties that open possibilities for opportunistic litigation and potential instrumentalisation of DB rights to enclose automatically generated data.[113] The straightforward abrogation of *sui generis* DB rights is not necessary for our purposes,[114] but the exclusion of IoT-generated data through a limited amendment of the DB Directive would fully address the concerns described above.[115]

---

[110] Draft Data Act (n 4).
[111] ibid recital 84.
[112] Estelle Derclaye and Martin Husovec, 'Why the Sui Generis Database Clause in the Data Act is Counter-Productive and How to Improve It?' (8 March 2022), draft paper available at <https://ssrn.com/abstract=4052390> accessed 20 September 2022.
[113] Guido Noto La Diega, 'Artificial Intelligence and Databases in the Age of Big Machine Data' (2018) 25 *AIDA* 93, 95; and Lionel Bently and Estelle Derclaye, 'Study in Support of the Evaluation of Directive 96/9/EC on the Legal Protection of Databases' (European Commission, 2018), available at <http://publications.europa.eu/resource/cellar/244f227a-597d-11e8-ab41-01aa75ed71a1.0001.01/DOC_1 > accessed 20 September 2022, 95.
[114] Some experts have advocated for the complete derogation of the sui generis right: Tarkowski and Vogelezang (n 34) 10; Valeria Falce, 'L'"insostenibile leggerezza" delle regole sulle banche dati nell'Unione dell'innovazione' (2018) 4(5) Diritto Industriale 377, 399; European Parliament (ITRE & IMCO Joint self-initiative report 'Towards a Digital Single Market Act' of 21 December 2015 (2015/2147(INI)), point 108. There are recent examples of the abolition of IP rights with no creativity requirements in Europe, such as the 2013 Dutch *geschriftenbescherming* (copyright protection for non-original writings). However, a complete abrogation could be problematic in terms of removing existing property rights and breaking with institutional inertia: see Martin Husovec, 'The Fundamental Right to Property and the Protection of Investment: How Difficult is it to Repeal New Intellectual Property Rights?' in C Geiger (ed), *Research Handbook on Intellectual Property and Investment Law* (Edward Elgar 2020).
[115] DB Directive (n 88), notably art 7(1).

C. The case for the trade secrets

The justification of trade secrets

TS are a creation of Anglo-American common law,[116] with roots in medieval guild ordinances in England.[117] Through TS, the legal order protects a piece of information that an enterprise values and treats as confidential against misappropriation. The EU legislated on TS only recently,[118] so the case law is still scarce. [119] The objective of the TS Directive was twofold: to harmonise the divergent framework among Member States[120] and to incentivise cross-border innovation and knowledge sharing.[121]

The TS Directive crystallises an international push to provide protection for TS in the last decade,[122] and enshrines the Anglo-American conception,[123] with perhaps a slightly

[116] Mark Lemley, 'The Surprising Virtues of Treating Trade Secrets as IP Rights' (2008) 61 Stanford Law Review 311, 317. The origins can be traced back to 1817 in England (*Newberry v James* (1817) 35 ER 1011, 1013) and 1837 in the US (*Vickey v Welch* (1837) 36 Mass 523, 527). The judicial recognition of TS takes place in the context of industrialisation and the opening of a certain labour mobility: see Margo EK Reder and Christine Neylon O'Brien, 'Managing the Risk of Trade Secret Loss Due to Job Mobility in an Innovation Economy with the Theory of Inevitable Disclosure' (2012) 12 Journal of High Technology Law 373, 386.
[117] Sean Bottomley, 'The Origins of Trade Secrecy Law in England 1600–1851', (2017) 38(3) The Journal of Legal History 254.
[118] Parliament and Council Directive (EU) 2016/943 of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ 2016 L157/1. The deadline for implementation expired on 9 June 2018.
[119] Jens Schovsbo, Timo Minssen and Thomas Riis, 'An Appraisal of the EU Directive on Trade Secrets' in Jens Schovsbo et al (n 36) 3.
[120] Hence, the legal basis is art 114 of the Treaty on the Functioning of the European Union (Consolidated version) OJ 626/47 (TFEU) (the approximation of legislation affecting the internal market) and not art 118 of the TFEU (unitary regime of protection for IP rights). Member States were more or less evenly divided among those with specific civil law provisions on TS (Bulgaria, Czech Republic, Denmark, Estonia, Germany, Italy, Lithuania, Poland, Portugal, Slovakia, Slovenia, Spain, and Sweden) and those without, the later relying instead on a general clause of prohibition of unfair competition, tort law, contract law, labour law or criminal law. See Davide Arcidiacono, 'The Trade Secrets Directive in the International Legal Framework' (2016) 1(3) European Papers 1073, 1077.
[121] TS Directive (n 117), recitals 1–4 and 8.
[122] The 1995 TRIPS agreement includes the protection of undisclosed information in Section 7 of Part II. This effort was led by the US: see Sharon Sandeen, 'Through the Looking Glass: Trade Secret Harmonization as a Reflection of US law' in Schovsbo et al (n 36) 41; Pistor (n 21) 121. Paradoxically, this push is now successful at the time when US scholarship begins to question the underpinnings of TS in the digital world, as we will see below in section 3.C.
[123] The TS directive contains the same characteristics as in the US, namely, a) TS are pieces of information which have commercial value because they are secret and have been subject to reasonable efforts to be kept secret (art 2(1) of TS Directive, in a similar vein to Section 1(4) of the US Defence of Trade Secrets Act (DTSA) and art 39(2) TRIPS agreement); b) TS are not registered and are only considered as defensive rights; that is, their holder can be entitled to legal remedies against misappropriation, but without a formal

different approach to its legal nature. That is, under EU law, TS do not constitute, strictly speaking, IP rights.[124] In the US, although the question remains debatable for lack of a clear indication in the Defence of Trade Secrets Act (DTSA),[125] they are treated as IP for practical purposes.[126]

I claim that the justification of TS does not hold in the data economy. To prove this, it is of particular interest to look at the US doctrine where the initial justification was more widely elaborated and its evolution has been subject to a richer scholarly scrutiny. With the US Supreme Court arguing that the justification of TS lays on 'the maintenance of standards of commercial ethics and the encouragement of invention',[127] the tenets for the justification of TS (and their expansion) revolve around:

1) Utilitarian arguments:

a) TS provide incentives to innovate;[128] and

property right (art 12 of TS Directive and Section 2 of the DTSA); c) TS can be acquired legally through independent discovery or reverse engineering (art 3 TS Directive and Section 1(1) DTSA). There are some small differences in the protection that the EU Directive and the US DTSA provide (in the EU, protection is in certain instances broader): for a detailed account see Sandeen (n 121) 60.

[124] Recital 2 speaks of TS as 'protecting a wide range of know-how and business information, whether as a complement or as an alternative to intellectual property rights'. See further, Tanya Aplin, 'A Critical Evaluation of the Proposed Trade Secrets Directive' (2014) 4 Intellectual Property Quarterly 257, 264. Moreover, most Member States did not consider TS as IP before the TS directive: see Tanya Aplin, 'Right to Property and Trade Secrets' in C Geiger (ed), *Research Handbook on Human Rights and Intellectual Property* (Edward Elgar 2015), 422. This is relevant for competition law, as indicated in Commission Statement concerning art 2 of Council Directive 2004/48/EC on the enforcement of intellectual property rights [2004] OJ 2004 L94/37; Henrik Udsen, Jens Schovsbo and Berdien van der Donk, 'Trade Secrets Law as Part of Information Law' in Schovsbo et al (n 36) 41. However, this nuance has not made TS less of an obstacle in accessing data according to the EU case law (see n 191).

[125] Charles T Graves, 'Trade Secrets as Property: Theory and Consequences' (2007) 15 Journal of Intellectual Property Law 39 (2007). In favour of the IP nature, Richard A Epstein, 'The Constitutional Protection of Trade Secrets under the Takings Clause' (2004) 71 University of Chicago Law Review 57, invokes the taking clause of the US Constitution, applicable both to TS and property; against it, Pamela Samuelson, 'Privacy as Intellectual Property' (2000) 52 Stanford Law Review 1125, notes that liability only stems from a wrongful acquisition. There are also intermediate theories, such as Lionel Bently, 'Trade Secrets: "Intellectual Property" But Not "Property"?' in J Griffiths and H Howe (eds), *Concepts of Property in Intellectual Property Law* (CUP 2013).

[126] Including competition law: see Harry First, 'Trade Secrets and Antitrust Law' in Rochelle Dreyfuss and Katherine Strandburg (eds), *The Law and Theory of Trade Secrecy: A Handbook of Contemporary Research* (Edward Elgar 2011) 1.

[127] *Kewanee Oil Co. v. Bicron Corp*, 416 *US* 470 (1974).

[128] Jon Chally, 'The Law of Trade Secrets: Toward a More Efficient Approach' (2004) 57 *Vandebilt Law Review* 1269, 1270; Lemley (n 115) 341. Among defenders of TS, the innovation argument is downplayed by

b) avoid more costly measures to protect their secrets from disclosure,[129] while enabling licensing and increasing transparency.[130]

2) Moral arguments:

a) TS help to maintain ethical standards for the marketplace; [131] and

b) preserve confidential relationships,[132] in particular regarding former employees.[133]

3) Limitations: the legal order recognises TS under certain conditions and trade-offs such as the possibility that reverse engineering and independent discovery permit the lawful acquisition of a TS.[134]

TS have expanded beyond their original contours.[135] This enlargement has triggered a doctrinal backlash questioning whether the original justification is still respected.[136]

---

some, like Michael Risch, 'Why Do We Have Trade Secrets?' (2007) 11 Marquette Intellectual Property Law Review 1, 26.

[129] Douglas Lichtman, 'Property Rights on the Frontier: How the Law Responds to Self-Help' (2005) 1 Journal of Law, Economics and Policy 215, 232; Risch (n 127) 43.

[130] Lemley (n 115) 342–45.

[131] Kurt M Saunders, 'The Law and Ethics of Trade Secrets: A Case Study' (2006) 42 California West Law Review 209; Lemley (n 115) 313.

[132] The main element of TS was originally a breach of trust, that is, a relational wrong rather than the violation of an exclusive right, which explains that protection covers 'misappropriation' and not other possibilities of 'appropriation': see Amy Kapczynski, 'The Public History of Trade Secrets' (2022) 55 University of California, Davis 1367, 1388.

[133] *CVD Inc v Raytheon Co*, 769 F 2d 842 (1st Cir 1985). See also Steven Wilf, 'Trade Secrets, Property, and Social Relations' (2002) 34 Connecticut Law Review 787, 794.

[134] *Chicago Lock Co v Fanberg* 676 F 2d 400 (9th Cir 1982); Risch (note 127) 53.

[135] Kapczynski (n 131) 1391 argues that there emerged in the 1970s a neoliberal justification of TS based on efficiency, but this was different from the original interpretation, confined to the 'theft' of secrets by former employees. Today, judges have recognised TS in customer lists, pricing information, business plans or marketing data. See Jeanne Fromer, 'A Legal Tangle of Secrets and Disclosures in Trade: Tabor v Hoffman and Beyond' in Rochelle Cooper Dreyfuss and Jane C Ginsburg (eds), *Intellectual Property at the Edge: The Contested Contours of IP* (CUP 2014) 271.

[136] Robert Bone, "A New Look at Trade Secrets Law: Doctrine in Search of Justification" (1998) 86 California Law Review 541 has been at the centre stage of a doctrinal struggle in the early 21st century. In a later publication (Robert Bone, 'The (Still) Shaky Foundations of Trade Secret Law' (2014) Texas Law Review 1803), he challenged the attempts to debunk his critique.

Some observers posit that the expansion has broken the internal balance of interests upon which TS protection was constructed,[137] an expansion that other IP rights cannot attain given their more explicit statutory regulation.[138] They argue that TS are used increasingly as weapons for concealment,[139] which is unjustified in situations where there are no innovation incentives,[140] no lower costs of transaction,[141] or, ultimately, where other societal interests are more worthy of protection.[142]

Against this backdrop, it is relevant to revisit the (already questioned) justification for TS protection in the context of the data economy.

Trade secrets in the data economy; a reassessment of the justification

The specificities of the data economy exacerbate and add elements to the arguments against the expansion of TS:

1) Utilitarian arguments:

[137] Michael P Simpson, 'The Future of Innovation: Trade Secrets, Property Rights, and Protectionism: An Age-Old Tale' (2005) 70 Brooklyn Law Review 1121, 1122 ('developments in the area of trade secret law have swung the pendulum too far in the direction of industry').
[138] David Levine, 'Secrecy and Unaccountability: Trade Secrets in Our Public Infrastructure' (2007) 59 Florida Law Review 135, 151.
[139] Charles Tait Graves and Sonia K Kaityal, 'From Trade Secrecy to Seclusion' (2021) 109 Georgetown Law Review 1337, 1402 ('private actors have successfully pushed its traditional market-competitive boundaries to turn TS into a tool for open-ended concealment').
[140] First (n 125) 5 argues that TS concern appropriability, not innovation, as they protect regardless of the innovative character of the secret. See also Derek Bambauer, 'Secrecy is Dead – Long Live Trade Secrets' (2016) 93 Denver University Law Review 833.
[141] Michael Burstein, 'Exchanging Information Without Intellectual Property' (2012) 91 Texas Law Review 227 argues that the old Arrow's 'disclosure paradox' (contracting over information is difficult because if the information is revealed to the buyer to evaluate the prospective transaction he will no longer need to pay for it) is not a theoretically nor empirically justified assumption.
[142] These interests include: access to government information (David S Levine, 'The People's Trade Secrets' (2011) 18 Michigan Telecommunications and Technology Law Review 61); transparency on authorisation of drugs (Christopher Morten and Amy Kapczynski, 'The Big Data Regulator, Rebooted: Why and How the FDA Can and Should Disclose Confidential Data on Prescription Drugs and Vaccines' (2021) 109(2) California Law Review 493); or overcoming opacity in the digital realm (Sylvia Lu, 'Algorithmic Opacity, Private Accountability, and Corporate Social Disclosure in the Age of Artificial Intelligence' (2021) 23 Vanderbilt Law Review 99).

a) As to the incentives to innovate: setting up the means to acquire big data can be inventive and expensive, [143] and this could justify a patent over a device, but not TS protecting the very data collected. In many of the business models found frequently in the data economy, TS are over compensatory with no solid incentive to innovation,[144] such as those that extract data from individuals through data-for-access arrangements for later consumer profiling [145] or those where the automatic generation is a mere by-product of a system set up with other primary goals.[146] In addition to this lack of incentives, follow-on innovation is severely compromised when TS cover automatically generated data,[147] an effect that compounds by virtue of the general purpose of data (as it is possible to use for purposes different from the one intended by the data collector), and knowing that productivity in the digital economy is shaped by the action of information upon information.[148]

A final consideration in this context. The IP legal framework has distributive effects among different innovative activities. Tilting the direction of innovation might be prejudicial to some of the operators in the market, possibly to the advantage of others, but this does not necessarily entail 'stifling' innovation overall.[149] A regulatory approach of targeted softening of TS protection could be intended to benefit the type of innovative activity deemed the most socially desirable.[150]

[143] Jeanne Fromer, 'Machines as the New Oompa-Loompas: Trade Secrecy, the Cloud, Machine Learning, and Automation' (2019) 94 New York University Law Review 706, 721 illustrates this with the example of Facebook, which cultivates a huge system for the capture of data about its users.
[144] Guido Noto La Diega and Cristiana Sappa, 'The Internet of Things at the Intersection of Data Protection and Trade Secrets: Non-Conventional Paths to Counter Data Appropriation and Empower Consumers' (2020) 3 Revue européenne de droit de la consommation 419; Lev-Aretz and Strandburg (n 79) 293.
[145] Gianclaudio Malgieri, '"Ownership" of Customer (Big) Data in the European Union: Quasi-Property As Comparative Solution?' (2016) 20 Journal Internet Law 3. See also the examples referred to in section 3.C.
[146] This is the case with many of the so-called core platform services as elaborated in the impact assessment for the DMA proposal (SWD (2020) 364 final) 39–45.
[147]Lev-Aretz and Strandburg (n 79), 266.
[148] Manuel Castells, *The Rise of the Network Society* (Wiley 2010) 19.
[149] Lev-Aretz and Strandburg (n 80) 17.
[150] This is particularly so in the digital space, where the incentives would be in the stimulus of the exchange of information rather than in barriers to it: see Orly Lobel, *Talent Wants to be Free: Why We Should Learn to Love Leaks, Raids, and Free Riding* (2013 YUL) 98.

b) As to the facilitation of the licensing of secret information: given that data is susceptible to encryption and different technical means of supply,[151] TS would not make the contracting over data easier or more transparent. Indeed, secrecy can be maintained through technical means which are increasingly robust and inexpensive, which disproves the justification of TS as a way to avoid overtly costly systems of protection. This is further underpinned by the fact that, in the EU context, new laws are introducing secrecy obligations on data transactions irrespective of TS.[152]

2) Moral arguments:

a) As to the ethical standards for the marketplace:[153] the arguments gravitate around the unfair advantage that third parties (free-riders) would have over the TS holders. In the data economy, the free-riding problem does not exist or has very different contours;[154] for example, in the intermediary platform ecosystem the capture of customers' and sellers' data to generate a market of promises of predictability to advertisers actually inverses the situation and puts the data holder in the position of free-rider as exploiter of an imbalance of informational power. More generally, the free-riding situations that underlie the justification for IP (and TS) protection are based on the assumption that 'copying' something creative or innovative is possible. This assumption fails largely in the context of data. Additionally, new data laws in the EU are incorporating fair compensation obligations for data generators.[155]

b) As to the preservation of confidential relationships: the role workers' mobility on the data itself is minimal, as machine-mediated processes can incorporate technical tools to ensure secrecy,[156] on top of which new data laws are creating

[151] See above in section 2.A.
[152] As an example, see art 11 of the Draft Data Act (n 4) on provisions against unauthorised use or disclosure of data.
[153] Saunders (n 130) 209.
[154] Lev-Aretz and Strandburg (n 79) 293.
[155] Draft Data Act, (n 4), art 9.
[156] Fromer (n 142) 721.

further guarantees of non-disclosure.[157] Data-generating machines often enjoy IP protection. Offering an additional shield to the data via TS would be socially problematic in many circumstances, in particular when the exclusivity on such data pre-empts potential competition.[158]

3) Limitations, such as reverse engineering or independent discovery, are simply not possible or relevant. A piece of information kept secret can be independently discovered; it is difficult to imagine this for data *per se*.[159] Data holders can exclude reverse engineering of their TS, be it with technical means (shielding the valuable parts of software derived through machine learning),[160] or with legal means (invoking a possible patent over their data-generating technology).[161] Limitations play a big role in the justification of IP and TS. The protection of a patent or copyright ultimately falls after a period of time fixed in the law. In the case of TS, the lack of explicit term is justified by the existence of legal means of acquisition (reverse engineering or independent discovery). Without these limitations, the nominal 'indeterminate' term of protection turns into an actual 'indefinite' term of protection.

Added to those considerations, there are significant barriers for access peculiar to the data economy, the most important of which would be network effects. This means that the more users are in a certain network (such as social media), the more difficult it is for later entrants to offer attractive alternatives,[162] thus consolidating the model of data accumulation in exchange for access,[163] which presumably can be further insulated by TS.

[157] Draft Data Act (n 4) art 11. Another example: a Parliament and Council Directive (EU) 2019/944 on common rules for the internal market for electricity [2019] OJ 2019 L158), art 41.
[158] Brenda M Simon and Ted Sichelman, 'Data-Generating Patents' (2017) 111 Northwestern University Law Review 377, 3381.
[159] See above in section 2.A on the disambiguation between data and information.
[160] Amanda Levendowski, 'How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem' (2018) 93 Washington Law Review 579, 590.
[161] Simon and Sichelman, (n 157) 408 (when a data-generating invention is patented, would-be competitors are effectively foreclosed from reverse engineering for a twenty-year exclusivity period that the patent awards).
[162] Lev-Aretz and Strandburg (note 80) 16.
[163] Cohen (n 38) 154.

Finally, there are sectors in which the legislator might deem that these elements (lack of innovation, free-riding by data holders, technical opacity) are exacerbated. These situations are those where, as we will see in the following section, the legislator is enacting data access mandates. The justification for TS in these sectors is still more intensely undermined.

In sum, the tenets of the doctrine of TS protection, already under serious doctrinal scrutiny at present, crumble in the context of the data economy.

Trade secrets defence in the data economy

The plummeting costs of information have put traditional IP instruments such as copyright and patents under stress,[164] at the same time that technology provides the means to keep robust 'secrets'.[165] In this context, data-driven industries have turned to TS as a preferred method of protection in addition to what technical means can attain.[166] TS might lack some of the apparently stronger legal teeth of patents (or to a less extent copyright), but they also do not have the limitations (in time or uses), conditions (inventiveness or creativity), trade-offs (such as disclosure) and exceptions (such as fair use) that characterise the classical IP instruments.[167]

In view of the increasing value of data, digital industries are recurring to TS for their litigation strategies.[168] Examples of the deployment of a TS defence with respect to data

[164] Bambauer (n 139) 841.
[165] Fromer (n 142) 724.
[166] Laura Palk and Krishnamurty Muralidhar, 'A Free Ride: Data Brokers' Rent-Seeking Behavior and the Future of Data Inequality' (2018) 20 Vanderbilt Journal of Entertainment and Technology Law, 779, 783; Sonia Katyal, 'The Paradox of Source Code Secrecy' (2019) 104 Cornell Law Review 1183; Peter S Menell, 'Tailoring a Public Policy Exception to Trade Secret Protection' (2017) 105 California Law Review 1, 3; Gianclaudio Malgieri, '"Ownership" of Customer (Big) Data in the European Union: Quasi-Property As Comparative Solution?' (2016) 20 Journal of Internet Law 3. This reliance on TS protection already started decades ago for software developers: see Mark Lemley and David O'Brien, 'Encouraging Software Reuse', (1997) 49 Stanford Law Review 255, 258.
[167] Deepa Vardarajan, 'Trade Secret Fair Use' (2014) 83 Fordham Law Review 1401, 1405.
[168] Daniel Gervais, 'Exploring the Interfaces Between Big Data and Intellectual Property Law' (2019) 10(1) Journal of Intellectual Property, Information Technology and E-Commerce Law 3; Teresa Scassa, 'Data Ownership' (2018) 187 CIGI Papers Working Paper No 2018-26 1; Matt Malone, 'Trade Secrets, Big Data, and the Future of Public Interest Litigation Over Artificial Intelligence in Canada' (2020) 35 Canadian Intellectual Property Review 6, 6; Cohen (n 37) 55.

abound.[169] When access was requested to Uber and Lyft's ZIP-code-indexed ride data in the city of Seattle under suspicion of racial bias, the ride-hailing apps claimed TS protection.[170] When the New York City passed, in August 2021, a bill requiring food delivery apps to share with restaurants data on their customers,[171] with the intention of equalising the information asymmetries of the market, Uber Eats, Grubhub and DoorDash filed a suit advancing the argument on the illegitimate expropriation of their TS (not on their software, but on 'their' data).[172] When the dating app Tinder was pressed for transparency as to the desirability rankings of profiles, it asserted TS defence.[173] When Myriad Genetics, a provider of genetic testing services in the context of cancer, was denied patents on its technology, it turned to TS for its large patient data.[174]

While judges can correct the most egregious cases of over claim, the phenomenon is undeniable. It is startling that cutting-edge companies rely on an instrument born in the realm of guild protectionism rather than the instruments of innovative free markets like patents.[175]

A view from the EU on trade secrets in the data economy

[169] Cristiana Sappa 'What Does Trade Secrecy Have To Do with the Interconnection-Based Paradigm of the Internet of Things?' (2018) 40(8) European Intellectual Property Review 518) describes the situations in which big data is processed to generate correlations between persons and their preferences.
[170] The case was eventually lost by Uber and Lyft, but not because their data was not considered TS (the court confirmed that such data constitutes TS), but because the data had been transferred to the city of Seattle as part of a deal and the Court found no reason to exclude such information from the scope of the State's Public Records Act. *Lyft Inc and Rasier LLC v City of Seattle*, 94026-6 (WN 2018).
[171] New York City Local Law 2021/090 to amend the administrative code of the city of New York, in relation to data on orders placed through third-party food delivery services <https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4951001&GUID=4CB11989-5925-418B-9627-B2AED230D67F&Options=&Search> accessed 20 September 2022.
[172] The suit by DoorDash states that 'by forcing DoorDash to disclose that TS to restaurants, the ordinance eliminates DoorDash's central property right in the TS – the right to exclusive use, and the right to exclusive use is the reason the trade secret has economic value.' See "DoorDash sues New York City over new data sharing law' *CNBC*, 15 September 2021 <https://www.cnbc.com/2021/09/15/doordash-sues-new-york-city-over-new-data-sharing-law.html> accessed 20 September 2022.
[173] Malone (n 167) 6.
[174] Simon and Sichelman (n 157) 378.
[175] Pistor (note 21) 128.

In the EU, the alarms are not ringing sufficiently regarding the possible interferences of TS in the data economy. The Commission argued in 2017[176] that:

> it is doubtful that individual data generated by interconnected machines and devices could be regarded as TS in the sense of this Directive, mostly because of its lack of commercial value as individual data; however, combination of data (datasets) can be trade secrets under this Directive if all the criteria are met.

The proposal for a Data Act explicitly maintains TS as a legitimate limit to data sharing obligations,[177] which contrasts starkly with the approach followed regarding DB rights,[178] despite both being part of the reflections in the ESD on the interference of IP rights in the data economy.[179] This could change during the legislative process, but it has so far not emerged as a point of contention.[180]

The doubts cast over the application of TS linked to the condition of commercial value are surprising. The TS Directive lays down a wide definition of 'commercial value' in its recital 14:

> Such know-how or information should be considered to have a commercial value (whether actual or potential), for example, where its unlawful acquisition, use or disclosure is likely to harm the interests of the person lawfully controlling it, in that it undermines that person's scientific and technical potential, business or financial interests, strategic positions or ability to compete.

In view of this large interpretation, many datasets created in the current data economy would fall under the definition of TS (provided that they are kept secret), even if by-

[176] Commission SWD (n 70) 20.
[177] Draft Data Act (n 4) art 8(6).
[178] See section 3.B.
[179] See section 2.B..
[180] The European Parliament's self-initiative report (n 51) adopted in reaction to the Commission's ESD 'urges the Commission to incentivise businesses to exchange their data, whether original, derived or co-generated, possibly through a reward system and other incentives, while respecting TS, sensitive data and IPR'.

products of another commercial activity.[181] For example, geo-location data, so often compiled indiscriminately by smartphone apps, definitely has such commercial value, as AI technologies can manage such data efficiently. Indeed, markets of data are emerging in different fields, a definite evidence of its 'commercial value'.[182]

Moreover, unlike *sui generis* DB protection, TS can apply to both data analysis techniques and datasets per se,[183] and within those, to both the semantic and syntactic level of data.[184] This has led some observers to conclude that TS operate in the data economy as 'legal intensifier[s] of factual exclusivity',[185] interfering severely in the effort to promote data sharing among private actors and with the public sector. Also, unlike the DB protection, the examples of TS litigation by data holders are proliferating and cannot be ignored.

D. Mandatory data access

There is a growing interest in the legal possibilities of mandating access to data.[186] There are two ways to attain this: firstly, as part of the enforcement of competition policy, when a refusal to share data is considered an abuse of dominant position, and secondly, as a result of explicit sector-specific statutory mandates.[187]

Competition law

[181] The fact that data is a by-product of a principal activity (such as selling goods) does not mean that it is at no cost for the operator, as it might decide to invest in its processes to improve the collection of data. See Alexandre De Streel, 'Big Data and Market Power' in Damien Gerard, Bernd Meyring, Eric Morgan de Rivery (eds), *Dynamic Markets, Dynamic Competition and Dynamic Enforcement: The Impact of the Digital Revolution and Globalisation on Competition Law Enforcement in Europe* (Bruylant 2018) 103.
[182] Mayer-Schönberger and Cukier (n 14) 90.
[183] Aplin (n 84) 67.
[184] Drexl (n 34).
[185] Zech (n 4) 26.
[186] Beata Mäihäniemi, *Competition Law and Big Data: Imposing Access to Information in Digital Markets* (Edward Elgar 2020); Inge Graef, *EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility* (Kluwer 2016).
[187] Depending on who is entitled to the access, we could also distinguish by whether the beneficiary is the agency or regulator or another private operator (a user or competitor). While this second criterion might lead to a different threshold at which access is granted (presumably lower when the beneficiary is a public entity), both types of access reveal the same concern for our purposes: whether TS protection is a viable defence about the obligation to share data.

While the idea of stipulating mandatory access to data in the law is relatively new, there already exists interesting case law in the field of competition law.[188] Some scholars have tried to articulate a doctrine of 'essential facilities' for the data economy,[189] but these attempts have so far not fructified. Be that as it may, the EU has been more open to the idea of access to data through competition law instruments than the US.[190]

A look into the EU case law is illuminating. In the *Microsoft* case[191] the General Court looked under the optic of 'abuse of dominant position' to an operator's refusal to share 'interoperability information' with a competitor. On an aside, let us clarify that this information would amount to TS, but the Court does not distinguish it for the purposes of their analysis from other IP rights.[192] The Court concludes that such refusal to license 'IP rights' can, under certain circumstances, constitute an abuse of dominant position.[193] This is not a foregone conclusion, for the Court relies on the 'exceptional circumstances' doctrine developed in previous copyright cases *Magill* ad *IMS Health*.[194] Moreover, the consideration that Microsoft was almost holding a position of monopoly in the market weighted significantly in the Court's decision, making these considerations hard to apply exactly to other cases.[195] Hence, although TS can be overridden by considerations of competition policy, this is conditional upon certain extraordinary circumstances equal to those applicable to traditional IP holders.

---

[188] Observers had detected that the issue of access to data on online platforms was likely to attract scrutiny of competition authorities and courts more than a decade ago. See Inge Graef, Sih Yuliana Wahyuningtyas and Peggy Valcke (2015) 'Assessing Data Access Issues in Online Platforms' 39(5) Telecommunications Policy 375, 379.

[189] Inge Graef, 'Rethinking the Essential Facilities Doctrine for the EU Digital Economy' (2019) 53(1) Revue juridique Thémis de l'Université de Montréal 33, lays down some proposals for reviving the 'essential facilities' doctrine; Nikolas Guggenberger, 'Essential Platforms' (2021) 24 Stanford Technology Law Review 237.

[190] Both the US antitrust authorities and the courts have been reluctant to force platforms to give access to their data on the basis of US antitrust law. See Graef, Wahyuningtyas and Valcke (n 187) 387.

[191] Case T-201/04 *Microsoft v Commission* (ECLI:EU:T:2007:289).

[192] ibid 289. This consideration underpins the idea that, despite a formally different legal nature, TS are treated as IP rights.

[193] ibid 690.

[194] Joined Cases C-241/91 P and 242/91 P *Magill* (ECLI:EU:C:1995:98) and Case C-418/01 *IMS Health GmbH* (ECLI:EU:C:2004:257). The abuse of dominant policy usually translates into the exclusion of competitors from a market, often with a burden to innovation: see Steven D Anderman and Hedvig Schmidt, 'EC Competition Policy and IPRs' in Steven D Anderman (ed), *The Interface Between Intellectual Property Rights and Competition Policy* (CUP 2007) 38.

[195] Pierre Larouche, 'The European Microsoft Case at the Crossroads of Competition Policy and Innovation' (2008) 75(3) Antitrust Law Journal 601, 628.

Some national competition authorities have found attempts to exclude others from data – such as customer lists – abusive. Both in France (*GDF Suez*) and Belgium (*Nationale Loterij*) the competition authorities fined companies for using customers lists acquired during their time as legal monopolies for the launch of new products, to the detriment of their competitors.[196] However, this does not amount to a direct mandate to share TS with competitors and it applies to the very peculiar circumstances of former legal monopolies.

More recently, the Commission has opened cases that entail an 'intrusion' in TS protection for the enforcement of competition policy as regards data. In the *Google Shopping* case the Commission fined Google for relegating competitors' shopping services in its search results. The Commission mandated Google to transmit its parameters (constituting TS) for the purposes of the investigation, although not to the competitors or to the public. Google's application was subsequently dismissed, although the matter of a possible TS violation was not expressly invoked.[197] A similar investigation is being carried out into Amazon's use of its sellers' data.[198]

We can conclude that TS protection is treated, without much elaboration, as analogous to IP rights and has a case-by-case relationship of prevalence with competition policy. TS assertions can prevail over competition policy enforcement except when a high threshold of abuse is surpassed. In sum, there are many limits to what competition law can do in the context of the data economy.[199]

Statutory data sharing mandates

[196] De Streel (n 180) 104.
[197] Case T-612/17 *Google Shopping* (ECLI:EU:T:2021:763).
[198] European Commission Press Release 'Antitrust: Commission sends Statement of Objections to Amazon for the use of non-public independent seller data and opens second investigation into its e-commerce business practices' of 10 November 2020, <https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2077> accessed 20 September 2022.
[199] Jens-Uwe Franck and Martin Peitz, 'Market Definition and Market Power in the Platform Economy' (2019) CERRE Report; Shreeja Sen, 'Time for Upgrade: Why Competition Law Not Enough for the Platform Economy' blogpost of 20 February 2022 < https://botpopuli.net/time-for-upgrade-why-competition-law-is-not-enough-for-the-platform-economy/> accessed 20 September 2022.

The second pathway consists of enacting provisions with data-sharing mandates.[200] Although there are examples of such provisions dating back several years, such as the Open Telecommunications Directive,[201] the REACH Regulation[202] and the INSPIRE Directive,[203] most of these provisions in EU law are very recent. Examples include the Payment Services Directive,[204] the Motor Vehicles Regulation[205] and the Electricity Directive.[206] The ESD flags some other fields where more of such provisions could be enacted, such as health, finance, energy and agriculture.[207]

More interesting, for its horizontal scope, are the new data access rights laid down in the proposals for a DMA and DSA. The DMA[208] would introduce apparently powerful rights of access to data that platforms will be obliged to grant. Article 6, in particular in paragraphs 8 and 9, confirms a wide-ranging obligation for gatekeepers to give access and transfer data to their business users.[209] It is unclear whether such data access rights would

---

[200] Understood as access to data different to the requester's own data, enshrined in art 20 of the GDPR for personal information and in a the Parliament and Council Directive (EU) 2019/770 on certain aspects concerning contracts for the supply of digital content and digital services [2019] OJ 2019 L136, art 16(4).

[201] Art 6(3), Parliament and Council Directive 98/10/EC on the application of open network provision (ONP) to voice telephony and on universal service for telecommunications in a competitive environment [1998] OJ 1998 L101/24. Concretely, the Directive ensures access to directories of subscribers of telephone services. The Court of Justice clarified the scope of the right of access and the allocation of costs regarding such access in Case C-109/03 *KPN Telecom BV v OPTA* (ECLI:EU:C:2004:749).

[202] Regulation (EC) 1907/2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) and establishing a European Chemicals Agency [2006] OJ 2006 L396, art 27.

[203] Parliament and Council Directive 2007/2/EC establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) [2007] OJ 2007 L108, art 17.

[204] Parliament and Council Directive (EU) 2015/2366 on payment services in the internal market [2015] OJ 2015 L337, art 67.

[205] Parliament and Council Regulation (EU) 2018/858 on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles [2018] OJ 2018 L151, art 66.

[206] Parliament and Council Directive (EU) 2019/944 (n 156), art 40.

[207] European Strategy on Data (n 51), 6.

[208] Digital Markets Act (n 4).

[209] "8. The gatekeeper shall provide advertisers and publishers, as well as third parties authorised by advertisers and publishers, upon their request and free of charge, with access to the performance measuring tools of the gatekeeper and the data necessary for advertisers and publishers to carry out their own independent verification of the advertisements inventory, including aggregated and non-aggregated data. Such data shall be provided in a manner that enables advertisers and publishers to run their own verification and measurement tools to assess the performance of the core platform services provided for by the gatekeepers.

9. The gatekeeper shall provide end users and third parties authorised by an end user, at their request and free of charge, with effective portability of data provided by the end user or generated through the activity of the end user in the context of the use of the relevant core platform service, including by providing, free of charge, tools to facilitate the effective exercise of such data portability, and including by the provision of continuous and real-time access to such data."

prevail over TS or other IP protection. Looking at its sibling DSA,[210] it seems that TS or IP protection would remain fully in place. The weaker right of access (because only intended for access for vetted researchers) in Article 40(5)(b) of the DSA comes with the disclaimer that the data holder can demand

> to amend the request, where it considers that it is unable to give access to the data requested because ... giving access to the data will lead to significant vulnerabilities for the security of its service or the protection of confidential information, in particular trade secrets.[211]

Hence, the core of these access rights is undermined by the possibility given to the platforms to refuse such access based on TS concerns.[212]

This interpretation of the prevalence of the TS protection would be underpinned by the Open Data Directive,[213] which targets public sector information only and aims at making as much of it available for re-use as possible. Following its recital 28, and according to what it calls the principle 'as open as possible, as closed as necessary', the 'concerns in relation to privacy, protection of personal data, confidentiality, national security, legitimate commercial interests, such as trade secrets, and to intellectual property rights of third parties should be duly taken into account'. This translates into the crystal-clear exclusion of TS from the application of the Directive, as laid down in Article 2(1)(d); in contrast once again with the exclusion of DB claims in the same circumstances (Article

[210] Digital Services Act (n 66). The Council's general approach (ST 13203 2021 COR 1 - NOTE) adopted on 24 November 2021 does not propose changes to this provision. The first reading of the European Parliament adopted on 20 January 2022 (P9_TA(2022)0014) ahead of intersintitutional negotiations, introduced an amendment to delete the specific reference 'in particular trade secrets'. This deletion would not fundamentally alter the considerations above.

[211] Underpinned by recital 64 *in fine* ('All requirements for access to data under that framework should be proportionate and appropriately protect the rights and legitimate interests, including trade secrets and other confidential information, of the platform and any other parties concerned, including the recipients of the service.')

[212] Nazrin Huseinzade, 'Algorithm Transparency: How to Eat the Cake and Have It Too' (European Law Blog 27 January 2021) <https://europeanlawblog.eu/2021/01/27/algorithm-transparency-how-to-eat-the-cake-and-have-it-too/> accessed 20 September 2022; Paddy Leerssen, 'Platform Research Access in Article 31 of the Digital Services Act (Verfassungsblog 7 September 2021) <https://verfassungsblog.de/power-dsa-dma-14/> accessed 20 September 2022.

[213] Open Data Directive (n 60).

1(6)). The more recent draft Data Act is more straightforward and provides Article 8(6) that 'an obligation to make data available to a data recipient shall not oblige the disclosure of trade secrets', a similar approach to the already enacted Data Governance Act (Article 5).[214]

In conclusion, despite the growing concerns about data exclusivity by the biggest operators of the data economy, the efforts of the competition authorities and the legislator to secure data access in some sectors, the TS 'wall' remains high and unabated. The newest legislative proposals aiming to establish new data rights will not change the situation; on the contrary, the opening for a judicial interpretation in favour of mandatory access will be excluded in view of the provisions described above. Ultimately, these clauses are based on an overreaching interpretation of TS, since, as discussed before, TS do not protect against any type of 'appropriation', but only against some instances of 'misappropriation'.

Data access rights over trade secrets as a solution

While judges can correct perhaps the most manifest cases of over claim of TS on data, litigation over what constitutes legitimate TS protection in the digital world can be unpredictable. Coincidentally, the TS Directive has been enacted when the digital revolution was in full swing, while in more experienced jurisdictions such as the US, there is a series of serious interrogations as to the application of TS in their full extent to the digital world. In any event, in the context of access to information, the European case law in other sectors has been generous as to the extent of what a TS (or commercial confidentiality) carve-out entails.[215] We cannot conclude that a strict interpretation of the notion of TS can reasonably limit the risk of interference of TS with data-sharing obligations.

---

[214] See section 2.B.

[215] Emilia Korkea-Aho and Päivi Leino, in 'Who Owns the Information held by EU Agencies? Weed Killers, Commercially Sensitive Information and Transparent and Participatory Governance' (2017) 54 Common Market Law Review 1059, 1068, argue in the context of access to documents of EU agencies that transparency exceptions for TS have been interpreted so broadly that they speak of a 'general presumption of non-disclosure.'

Although the arguments developed before could underpin a more overarching negation of the justification of TS in the data economy, a general carve-out of TS for the data world would be unnecessarily wide for our purposes. It is also not necessary to revert to some of the scholarly proposals formulated in the US to avoid the overreach of TS, such as a TS-specific 'misuse doctrine' excluding protection for cases of over claim so socially valuable uses would prevail,[216] or statutory safe harbours,[217] decided by the legislator in specific cases.[218]

EU law offers a more surgical alternative, which could be more easily accepted in the context of the enactment of specific data access rights. It would consist of spelling out that TS defence does not prevail over specific data-sharing mandates and for the enforcement of competition law, in line with the considerations of the preceding section.[219]

IV. Conclusions

It might seem that data came to its prominence abruptly and recently, but it is not the case. Its emergence has been exponential but gradual. In this process, the law has been adapting to something distinct from other 'objects of the law', and thus the adaptation has not been entirely satisfactory. Simplistic analogies that treat data as a resource have cemented this 'adaptive' approach of the law. Confusion between the concept of data and information do not help to disentangle the legal challenges raised either. In this context, knowing that a notion as elementary for law as property is not applicable to something as

[216] This would mirror the doctrine of misuse developed by courts and applied to copyright law, which revolves around questions like whether the IPR holder is acting in an anticompetitive way that cannot be addressed by antitrust rule of reasons standard, whether the IPR holder licensing conditions are too restrictive of socially valuable acts, whether the matter can be better channelled through patent protection and whether the IPR holder is engaging in abusive claims. See Deepa Varadajan, 'The Uses of IP Misuse' (2019) 68 Emory Law Review 739, 779.
[217] Vardarajan (n 215) 1446.
[218] ibid. Examples include medical device pricing data: Annemarie Bridy, 'Trade Secret Prices and High-Tech Devices: How Medical Device Manufacturers Are Seeking to Sustain Profits by Propertizing Prices' (2009) 17 Texas Intellectual Property Law Review 187, 189, or data on companies contracting with the government in the context of elections; David Levine, 'The Impact of Trade Secrecy on Public Transparency' in Rochelle C Dreyfuss and Katherine J Strandburg (n 125); as regards TS held by public bodies, Christopher Morten, 'Publicizing Corporate Secrets' (2022) 171 University of Pennsylvania Law Review (forthcoming) or a public policy exception to TS for whistleblowing (which already exists under EU law), Menell (n 165) 3.
[219] This could be added to art 5 of the TS Directive (n 117), which already includes a list of exceptions.

central to the economy as data, the temptation is too evident to exploit some instruments within the loose margins of IP law to articulate similar entitlements.

This happens in a context where the EU has resolved to promote and, in some sectors, mandate data sharing. Enormous competitive advantages already exist for the leading data operators by the mere control of data infrastructure and network effects. This position should not be reinforced by leveraging existing IP rights to their advantage. I claim that TS or DB rights, especially where data access provisions exist, should not serve the purpose of erecting 'legal walls' around data and thus of undermining access rights. Such legal engineering goes beyond the intent of these two instruments, enacted before data took the prominent role it has today.

The possible interference of IP with data-sharing goals has been a blind spot (deliberate or not) for the legislator. There is a remedy to it through limited amendments: for DB protection, providing that automatically generated data does not fall within its scope; for TS, establishing a clear prevalence of mandatory access provisions and competition law enforcement over TS defence.

This piece has been a call to rethink certain elements of IP law in view of something so new in legal characterisation and so central to our society as data. This is just one of several challenges to the goal of data sharing and, in general, to the regulation of data. There are others, such as the role of infrastructure, of 'technical opacity' and that of personal data protection laws. This article is thus a limited contribution to a wider debate on how to create a legal framework for the data economy aligned with our societal goals.