

Please cite as Guido Noto La Diega and Estelle Derclaye, ‘Opening Up Big Data for Sustainability: What Role for Database Rights in the Fourth Industrial Revolution?’, in Ole-Andreas Rognstad, Taina Pihlajarinne, Jukka Mähönen (eds), *Promoting Sustainable Innovation and the Circular Economy: Legal and Economic Aspects* (Routledge 2024)

Opening Up Big Data for Sustainability: What Role for Database Rights in the Fourth Industrial Revolution?

Guido Noto La Diega and Estelle Derclaye*

Abstract

While scarcity of resources is the main traditional justification for proprietary approaches, the Fourth Industrial Revolution (4IR) heralds an age of abundance of data, which calls for a commons-inspired approach to data. Open data is no longer just the mantra of sparse communities of developers and designers. The EU has recently started to embrace open data in a number of recent legislative instruments and proposals to push public entities and citizens to open up their data for the common good. Liberating data from its enclosures is pivotal to achieving the UN Sustainable Development Goals (SDG), but the real guardians of big data – the private corporations that are the key decision-makers in the 4IR – are not doing enough to facilitate sharing and re-use of data in the public interest, including the pursuit of climate justice. The main instruments recently proposed to curb big tech power do not seem to embrace openness in a meaningful way. While there may be instances where Intellectual Property (IP) reasons may justify some limitations in access to and re-use of big data held by corporations, it is our view that, in general, IP should not be used to hinder re-use of data to pursue the SDGs. As governments and citizens are being asked to open up data, now is the time for private companies to open up as well. As the fight for sustainability has never been more pressing and as transparency rises as one of the pillars of global digital constitutionalism, private power must do better and more to allow access to and re-use of data for the common good. Against that backdrop, this paper aims to answer the following overarching research question: what does a sustainable legal framework for achieving sustainability goals look like in the 4IR? To answer it we will, first, illustrate the triple

* This chapter is the result of a genuinely collaborative effort between the authors. Guido Noto La Diega is primarily responsible for sections 1-4, Estelle Derclaye for section 5, both for section 6. The authors wish to thank the anonymous peer reviewers for their useful comments and suggestions. We also thank Ana Ramalho, Maria Lillà Montagnani, Martin Senftleben, and Séverine Dusollier for their comments at the 2022 ALAI and EPIP annual conferences.

meaning of ‘data sustainability.’ Second, we will critically assess whether the database right (or ‘sui generis right’) can play a role in opening up corporate big data. Third, we will imagine what a sustainable framework for sustainable data governance might look like. This focus is justified by the fact that the Database Directive, often accused of creating an unjustified monopoly on data, is in the process of being reformed by the latest addition to the EU data governance framework: the Data Act. A step in the right direction, the Data Act does not provide the required legal certainty in the machine data space, thus potentially frustrating its objective to leverage data for the green and digital transitions.

Key words: machine data, open data, data access, IP, EU Data Act, database rights, sui generis right, Fourth Industrial Revolution, data governance, connected products, product data, sustainability

1. Context and scope of study

In the Fourth Industrial Revolution (4IR), the growing fusion and interaction between technologies such as the Internet of Things (IoT), Artificial Intelligence (AI), blockchain, quantum computing, and nanotechnologies are rewriting the boundaries between the biological, the physical, and the digital.¹ As hundreds of zettabytes of data are being generated in the 4IR, it has become untenable to argue that the law needs to introduce incentives to produce more data.² Instead, as private corporations amass a wealth of data, we need incentives to open up enclosed data. Scarcity of resources is the main traditional justification for proprietary approaches.³ The 4IR heralds an age of abundance of data,⁴ which justifies – even calls for – a commons-inspired open data approach.⁵ This data is often

¹ See Klaus Schwab, *The Fourth Industrial Revolution* (Penguin 2017). 4IR is often confused with Industry 4.0, which is more focused on the industrial applications of the IoT (‘smart manufacturing’) see, eg, Christoph Winterhalter, ‘A New Revolution in the Making’ (2018) November-December ISOfocus 3.

² ‘Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2025’ (Statista, 2022) <<https://www.statista.com/statistics/871513/worldwide-data-created/>> accessed 12 January 2022.

³ Mark A Lemley, ‘IP in a World Without Scarcity’ (2015) 90 NYU Law Review 461.

⁴ See Andrea Ottolia and Cristiana Sappa, ‘A Topography of Data Commons: From Regulation to Private Dynamism’ [2021] *GRUR International* 1. The authors convincingly argue that a data commons approach is key to avoiding excessive protection in the data-driven society. Unlike the authors, we aim at stating that the commons is the best solution to data governance in the 4IR.

⁵ Data commons means not only open data but also collective practices such as co-creation of data protection solutions to rebalance power between data subjects and data controllers. See Janis Wong, Tristan Henderson,

referred to as ‘machine data’ but in this chapter we refer to ‘4IR data’ for a twofold reason. First, the phrase ‘machine data’ can be used to hide the value of this data, as if it were data merely produced by a machine and about a machine. Instead, this data is often produced by humans, related to humans (both personal and non-personal data), and is merely processed by machines. Second, while our considerations may have wider application, the scope of this chapter is the 4IR, as increasing blurring between biological, cyber, and physical systems demands closer scholarly scrutiny. So 4IR data means machine-generated data, whether personal or non-personal. However, in this chapter, for reasons of brevity, we focus solely on non-personal machine-generated data.

Open data is no longer just a mantra among sparse communities of developers and designers. It is being used increasingly in the Global South, from the ‘commoning’ of data to plan a smarter Rio de Janeiro to streamlining emergency responses in the Philippines.⁶ The EU has recently started to embrace open data in several recent legislative instruments and proposals to push public entities and citizens to open up their data for the common good. These in turn are epitomized by the new Open Data Directive⁷ and the Data Governance Act.⁸ The first of these creates the conditions for better access to and re-use of public sector information in the belief that openness can benefit society at large.⁹ The second of these invites citizens to embrace a philosophy of data altruism on the grounds that better data

and Kirstie Ball, ‘Data Protection for the Common Good: Developing a Framework for a Data Protection-Focused Data Commons’ (2022) 4 Data & Policy 1.

⁶ Fiona Smith, William Gerry and Emma Truswell, ‘Supporting Sustainable Development with Open Data’ (2015) Open Data Institute.

⁷ Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information [2019] OJ L 172/56. This instrument recast a 2003 directive that, even in its 2013 amendment, had not fully embraced a philosophy of openness while positively incentivizing the re-use of documents held by public sector bodies (see Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information [1003] OJ L 345/90). The rationale of the reform was ‘to bring the legislative framework *up to date with the advances in digital technologies* and to further stimulate digital innovation, especially with regard to *artificial intelligence*’ (Open Data Directive, Rec 3).

⁸ Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act) COM/2020/767 final.

⁹ Open Data Directive, Rec 8.

sharing can not only bring economic value, but also – and more significantly – ‘new ways of tackling societal challenges (for instance, climate change).’¹⁰

Liberating data from its enclosures is pivotal to achieving the UN Sustainable Development Goals (SDG). For example, the European Tracking Network uses open sensors (fish tags) and open data to integrate all aquatic animal tracking in Europe in one network; thanks to this shift to ‘a sustainable, efficient, and integrated pan-European biotelemetry network,’¹¹ we can better understand and protect aquatic animals, thus pursuing SDG 14 (‘Life below Water’). In 2020, in the context of the EU Green Deal, the European Commission laid the foundation for a Green Data Space, which will inter alia include reusable data services to monitor compliance with environment legislation and a common European data space for smart circular applications, a framework that will make ‘available the most relevant data for enabling circular value creation along supply chains.’¹² A European green data space would be a step in the right direction, but it is questionable whether it provides sufficient incentives for big 4IR to share their data. As we will argue in the next section, it is apparent that the real guardians of big data – the private corporations that are the key decision-makers in the 4IR – persist in not doing enough to facilitate the sharing and re-use of data in the public interest, including the pursuit of climate justice. The main instruments recently proposed to curb big tech power do not seem to embrace openness in any meaningful way. This is epitomized by the proposed AI Act.¹³ There, market surveillance authorities are granted access to the training, validation and testing datasets used by a provider of high-risk AI systems, and under certain conditions also to the relevant source code.¹⁴ This is a severely limited right to access, which does not allow re-use of this data and code for the common good, including for sustainability purposes. Crucially, this is consistent with the pro-proprietary stance of the AI Act summed up in the promise that:

(t)he increased *transparency obligations* will also not disproportionately affect the right to protection of *intellectual property* (...) since they will be limited only to the *minimum necessary* information for

¹⁰ Data Governance Act, Explanatory Memorandum, para 3. See also Rec 35.

¹¹ ‘European Tracking Network’ <<https://europeantrackingnetwork.org/en/about-us>> accessed 11 May 2022.

¹² Communication ‘A European strategy for data’ COM/2020/66 final [2].

¹³ Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts, COM/2021/206 final.

¹⁴ AI Act, Art 64. d

individuals to exercise their right to an *effective remedy* and to the necessary transparency towards supervision and *enforcement authorities*.¹⁵

While there may be instances where Intellectual Property (IP) related reasons may justify some limitations in access to and re-use of big data held by corporations – chiefly trade secrets and database *sui generis* right protection – it is our view that, in general, IP should not be used to hinder re-use of data to pursue the SDGs, at least as far as is possible to respect the balance found in IP laws between protection and access.

As Governments and citizens are being asked to open up data, now is the time for private companies to open up as well.¹⁶ As the fight for sustainability has never been more pressing and with the rise of transparency as one of the pillars of global digital constitutionalism,¹⁷ private power must do more and better to allow access to and re-use of data for the common good. With that in mind, this chapter aims to answer the following overarching research question: what does a sustainable legal framework for achievement of sustainability goals look like in the 4IR? First, we will illustrate the triple meaning of ‘data sustainability.’ In each of these meanings, data has to be open in order to maximize the chances that it be used to achieve the SDGs. Second, we will critically assess whether the database right (or ‘*sui generis* right’) can play a role in opening up corporate big data. Third, we will imagine what a sustainable framework for sustainable data governance might look like. This focus is justified by the fact that the Database Directive, often accused of creating an unjustified monopoly on data, is in the process of being reformed by the proposed Data Act, published by the Commission in February 2022, finalised in June 2023, but not yet published in the Official Journal.¹⁸ Trade secret protection is outside the scope of this paper, even if it is an important part of the debate, and should be further researched.¹⁹ Sharing of 4IR data with public

¹⁵ AI Act, Preamble, para 3.5.

¹⁶ As noted by the Expert Group on Business-to-Government Data Sharing, several barriers prevent scalable, responsible, and sustainable business-to-government data sharing in the public interest. See High-Level Expert Group on Business-to-Government Data Sharing, ‘Towards a European Strategy on Business-to-Government Data Sharing for the Public Interest’ (2020) Final report.

¹⁷ See, eg, Piotr Mikuli and Grzegorz Kuca (eds), *Accountability and the Law: Rights, Authority and Transparency of Public Power* (Routledge 2021).

¹⁸ Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (‘Draft Data Act’) COM(2022) 68 final. The text has been last amended by the Council on 7 July 2023 (OR. en), 2022/0047(COD) and we comment on this latest version here.

¹⁹ Regrettably, the Data Act leaves trade secrets on 4IR data largely unaffected. Future research should critically assess whether this nullifies the Act’s objective of opening up machine data for the common good.

authorities is also outside the scope of this paper, that is focused on the access and re-use of data by private IoT users.²⁰

2. Data sustainability: a multifaceted notion

Sustainability has traditionally been the domain of environmental law and policy.²¹ In recent times, as the world is waking up to the disastrously tangible consequences of climate change, legal scholars from all disciplines have been endeavouring to reflect on how their fields of expertise can contribute to achieving climate justice and, more generally, the SDGs.²² Experts in the laws of data are no exception.²³

Whilst the concept of ‘data sustainability’ has not yet been the object of systematic and holistic investigation, we posit that at least three meanings can be attached to it. In particular, data sustainability can refer to:

- (i) use of data to monitor progress in pursuit of the SDGs;
- (ii) development of data-driven technological solutions to make progress towards sustainability, including AI, the IoT, and the 4IR more generally;
- (iii) overcoming unsustainable data practices, that is, a vast range of abusive and extractive practices that 4IR companies put in place to the detriment of society.

The first meaning is the least problematic. Data sustainability in the sense of an SDG monitoring tool requires that data be open. Open data is pivotal to sustainability as a vital interest to provide critical information on natural resources, population demographics, public services, and government operations; thus informing ‘national priorities and help determine

²⁰ The starting point of future research endeavours should be Recital 63 of the Data Act whereby ‘[i]n case the *sui generis* database rights [...] apply in relation to the requested datasets, data holders should exercise their rights in a way that does not prevent the public sector body, the Commission, the European Central Bank or bodies from obtaining the data, or from sharing it, in accordance with this Regulation.’

²¹ MN Boeve and others (eds), *Environmental Law for Transitions to Sustainability* (Intersentia 2021).

²² The book this chapter contributes to is a prime example of this new approach to legal scholarship. See also Volker Mauerhofer, Daniela Rupo, and Lara Tarquinio (eds), *Sustainability and Law: General and Specific Aspects* (Springer International Publishing 2020).

²³ Luciano Floridi, ‘The Green and the Blue: Naïve Ideas to Improve Politics in a Mature Information Society’, *The 2018 yearbook of the digital ethics lab* (Springer 2019); Carla Barbosa and Andreia da Costa Andrade, ‘Biobanks and GDPR: A Look at the Portuguese Panorama’, *GDPR and Biobanking* (Springer, Cham 2021).

the most effective paths for action on national issues.’²⁴ A growing number of open data repositories, hubs, and databases are being made available to allow stakeholders and the general public to engage with the progress made so far in achieving the SDGs. These initiatives fall mainly under three categories: global SDG indicators, thematic and regional data resources, and progress reports.²⁵ Limiting ourselves to the first category, which is the most information-rich, in July 2017, the UN General Assembly adopted a resolution containing the global indicator framework for SDGs.²⁶ Thanks to the resulting 231 unique indicators it becomes easy to measure the progress made in the sustainability agenda. For example, an indicator of SDG 5 (‘Achieve gender equality and empower all women and girls’) is the ‘(p)roportion of individuals who own a mobile telephone, by sex.’²⁷ Whereas until recent time SDG-related data were siloed and hard to access, the UN have made progress by embracing openness as the way to ensure that decision makers around the world can implement the SDGs and monitor the relevant progress through open, timely, detailed, accessible, and easy-to-use data. This can be seen with clarity in the Open SDG Data Hub, which ‘promotes the exploration, analysis and use of authoritative SDG data sources for evidence-based decision-making and advocacy.’²⁸ Through the hub we can easily access and re-use all sorts of data, including geospatially referenced data. For example, we can assess the progress that Abu Dhabi has made with regard to SDG 11 (‘Sustainable cities and communities’) by looking at the proportion of urban solid waste regularly collected and with adequate final discharge (currently 17.49%). Such 4IR technologies and models as machine learning can then be used to understand patterns and interrelationships in SDG-related big data.

²⁴ Joel Gurin, Laura Manley, and Audrey Ariss, ‘Sustainable Development Goals and Open Data’ (*World Bank Digital Development*, 25 September 2015) <<https://blogs.worldbank.org/digital-development/sustainable-development-goals-and-open-data>> accessed 11 May 2022.

²⁵ This tripartition is taken from ‘Data Resources’ (*UN Stats*) <<https://unstats.un.org/sdgs/unct-toolkit/data-resources/>> accessed 13 January 2022.

²⁶ Resolution adopted by the General Assembly on Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development (A/RES/71/313), Annex.

²⁷ Global indicator framework adopted by the General Assembly in A/RES/71/313 (Annex), annual refinements contained in E/CN.3/2018/2 (Annex II), E/CN.3/2019/2 (Annex II), 2020 Comprehensive Review changes (Annex II) and annual refinements (Annex III) contained in E/CN.3/2020/2 and annual refinements (Annex) contained in E/CN.3/2021/2, target 5.b, 7.

²⁸ ‘Welcome to the Open SDG Data Hub’ (*UN Stats*) <<https://unstats-undesa.opendata.arcgis.com/>> accessed 13 January 2022.

This leads to the second meaning of ‘data sustainability’, which is interwoven with the first. Technological innovation and equitable technology adoption are pivotal to achieving the SDGs. For example, in order to achieve SDG 1 (‘No poverty’) a key target set by the UN is that all men and women have ‘ownership and control over (...) new technology.’²⁹ A 4IR example of this type is the Hub of All Things, a technology created from over £10m of UK research to enable individuals to own their own personal data server.³⁰ These 4IR technologies are vital for virtually all SDGs, from zero hunger,³¹ through quality education,³² to clean water,³³ affordable and clean energy,³⁴ decent work and economic growth,³⁵ industry, innovation and infrastructure,³⁶ sustainable consumption and production,³⁷ life below water,³⁸ revitalizing the Global Partnership for Sustainable Development,³⁹ and, as we have seen, gender equality.⁴⁰ Also thanks to the Open SDG Data Hub, the UN have started to reflect on how to use big data for sustainability and it is apparent that the 4IR will play a significant role in this direction. Indeed, in putting forward examples of how big data can be used to help achieve the SDGs, goal by goal, it becomes apparent that the 4IR is set to provide most of the data-driven solutions to make progress towards sustainability.⁴¹ For example, the UN sees mapping the movement of mobile phone users as a way to predict the spread of infectious diseases as necessary in order to pursue SDG 3 (‘Good health and wellbeing’).⁴² The list goes on and includes smart water pumps for clean water, smart meters for affordable energy, and GPS data to improve public transport.⁴³

²⁹ Global indicator framework (n 21), target 1.4.

³⁰ ‘Hub of All Things’ (HAT) <<https://www.hubofallthings.com/>> accessed 11 May 2022.

³¹ SDG 2, target 2.a.

³² SDG 4, targets 4.3, 4.4, 4.5, 4.b.

³³ SDG 6, target 6.a.

³⁴ SDG 7, targets 7.1, 7.a, 7.b,

³⁵ SDG 8, targets 8.2 and 8.a.

³⁶ SDG 9, targets 9.4, 9.5, 9.a, 9.b, and 9.c.

³⁷ SDG 12, target 12.a.

³⁸ SDG 14, target 14.a.

³⁹ SDG 17, targets 17.6, 17.8, 17.9, and 17.16.

⁴⁰ SDG 5, target 5.b.

⁴¹ ‘Big data for sustainable development’ (*United Nations*) <<https://www.un.org/en/global-issues/big-data-for-sustainable-development>> accessed 13 January 2022.

⁴² *ibid.*

⁴³ For an example of how new sharing-prone data practices can improve and personalize public transport by switching from a system of rigid bus routes and schedules to a fully dynamic, on-demand network, see ‘Via

Research abounds into how 4IR data can be used for sustainability purposes. A key reference is the special issue ‘Applications for Internet of Things’ of the journal *Sustainability* with important contributions ranging from the use of blockchain in healthcare⁴⁴ through autonomous sensor networks for smart agriculture⁴⁵ to smart crime prevention⁴⁶ and IoT data visualization for smart cities.⁴⁷ These articles, all deliberately published in open access, contribute to making a clear case in favour of openness as the most sustainability-friendly form of data governance. Crucially, a review of open government data practices and research shines a light on the ‘tendency to create sustainable open government data and smartness by employing artificial intelligence.’⁴⁸ While data-driven sustainability-oriented 4IR does not necessarily require data to be open, nevertheless arguably the re-usability, repurposing, auditability, transparency, and potential for sharing and collaboration that come with open data make it a much better candidate, compared to siloed, proprietary models. A couple of examples from the world of smart agriculture may help expand on this notion.⁴⁹ In the USA, private companies integrate private and public data to make progress towards nutritional health. For example, Nutritionix⁵⁰ is an app that uses open public data from the Department of Agriculture’s FoodData Central – a centralized gateway that makes comprehensive nutritional data on raw, processed and experimental food available for re-use – and combines it with private data (for example, restaurant menus) to help users with weight

Launches On-Demand Public Transit in Tel Aviv’ (Via, 14 April 2019) <<https://ridewithvia.com/news/via-launches-on-demand-public-transit-in-tel-aviv/>> accessed 11 May 2022.

⁴⁴ Muhammad Tahir and others, ‘A Lightweight Authentication and Authorization Framework for Blockchain-Enabled IoT Network in Health-Informatics’ (2020) 12 *Sustainability* 6960; Jesús Peral and others, ‘Using Visualization to Build Transparency in a Healthcare Blockchain Application’ (2020) 12 *Sustainability* 6768.

⁴⁵ Javier Rodríguez-Robles and others, ‘Autonomous Sensor Network for Rural Agriculture Environments, Low Cost, and Energy Self-Charge’ (2020) 12 *Sustainability* 5913.

⁴⁶ Mun-su Park and Hwansoo Lee, ‘Smart City Crime Prevention Services: The Incheon Free Economic Zone Case’ (2020) 12 *Sustainability* 5658.

⁴⁷ Ana Lavalley and others, ‘Improving Sustainability of Smart Cities through Visualization Techniques for Big Data from IoT Devices’ (2020) 12 *Sustainability* 5595.

⁴⁸ Yingying Gao, Marijn Janssen and Congcong Zhang, ‘Understanding the Evolution of Open Government Data Research: Towards Open Data Sustainability and Smartness’ [2021] *International Review of Administrative Sciences* 002085232110099.

⁴⁹ See Agata C Hidalgo, ‘5 Ways to Use Open Data in Agriculture’ (Opendatasoft, 25 February 2021) <<https://www.opendatasoft.com/en/blog/5-ways-to-use-open-data-in-agriculture>> accessed 11 May 2022.

⁵⁰ ‘Nutritionix - Largest Verified Nutrition Database’ (Nutritionix) <<https://www.nutritionix.com/>> accessed 11 May 2022.

concerns calculate recommended calory intake, track food consumption, and share their food log with an accredited dietitian. Another example is Japan's WAGRI,⁵¹ that aims at harmonizing data formats used in the farming sector to facilitate data sharing among farmers, agricultural businesses, equipment manufacturers, and software providers. Crucially, WAGRI also makes available data and software by private companies. Although many data-driven 4IR solutions are hybrid, it seems clear that openness plays an important role in achieving sustainability objectives.

The nexus between open data, the 4IR, and sustainability is further confirmed by an important set of guidelines that the World Economic Forum (WEF) has put together to make sure that 4IR companies do not sacrifice the development benefits of the IoT on the altar of commercial viability.⁵² Based on an analysis of over 640 IoT deployments worldwide, the WEF noted that 84% of them have at least the potential to address the SDGs.⁵³ To encourage prioritization of sustainability goals as part of the design of commercial IoT projects, a key guideline is to '(e)stablish early the data governance terms for ownership, privacy, usage and sharing.'⁵⁴ For the 4IR to express its sustainability potential, new collaboration models are needed; at its core, clear options in terms of use and sharing of data are necessary: they can go from sharing of complete datasets through filtering out of privacy-critical data to sharing of macro-level insights.⁵⁵ Whilst collaboration is key, new business and investment models are no less important. Against this backdrop, a relevant guideline asks private stakeholders to develop cross-industry solutions to unlock mutual benefits and enable new monetization models. Whereas we have become accustomed to business models founded on extreme data monetization, it is refreshing to see that the WEF is aware that transparency is pivotal to connections between previously unconnected industry and that, to this end, companies must open up data. Indeed, if 4IR enterprises are 'serious about innovation (they) should consider giving access to the data to other players while complying with data laws.'⁵⁶

The third meaning of 'data sustainability' is based on a perhaps less immediate link between the terms of the relation. The big players of the 4IR have been accused of a vast

⁵¹ 'WAGRI - Agriculture Data Platform' <<https://wagri.net/en-us/>> accessed 11 May 2022.

⁵² World Economic Forum, 'Internet of Things. Guidelines for Sustainability' (2018) Future of Digital Economy and Society System Initiative.

⁵³ *ibid* 3.

⁵⁴ *ibid* 4.

⁵⁵ *ibid* 11.

⁵⁶ *ibid* 13.

range of data abuses or unsustainable data practices, in violation – or, more often, in circumvention – of the laws on data governance, including IP, data protection, and competition. As noted in the proposed Data Act, data governance laws provide the basis for ‘sustainable and responsible data processing.’⁵⁷ Instances of unsustainable data practices are use of big data to accentuate inequalities, sale of patient data to private business, algorithmic bias, digital discrimination, predictive policing, automated enforcement, and what we can call ‘digital dispossession.’ As the latter is an understudied issue that sheds light on a more general problem with the 4IR – indeed, one that could hinder its potential to contribute to the sustainability agenda – we will dedicate the next section to it.

3. Digital dispossession or enclosure of data

Alongside economic and political power, technology-intensive industries have seen the rise of another, and possibly more insidious, form of power: data power. This ‘multifaceted power that may overlap with economic (market) power’ stemming from accumulation of user-generated and user-related data, entails the power to profile, to influence opinion formation, and to impose contracts that purport to justify unfair and opaque practices, including appropriation and re-use of personal as well as non-personal data: what we call ‘digital dispossession.’⁵⁸ Recently, the EU lawmaker has shown awareness of this phenomenon; notably, the proposed Data Act will provide new access rights on the grounds that IoT manufacturers and designers tend to have ‘*exclusive control over the use of data* generated by the use of a product or related service, which contributes to lock-in effects and hinders market entry for players offering aftermarket services.’⁵⁹ We would posit that digital dispossession in the 4IR has even wider and possibly more insidious consequences.

This concept is rooted in the social theory of ‘accumulation by dispossession’ developed by David Harvey.⁶⁰ The social theorist criticized Marx⁶¹ and Rosa Luxemburg⁶²

⁵⁷ Draft Data Act, Rec 7.

⁵⁸ The aspects related to profiling and opinion formation have been dealt with by Orla Lynskey, ‘Grappling with “Data Power”: Normative Nudges from Data Protection and Privacy’ (2019) 20 *Theoretical Inquiries in Law* 189. The latter aspect in 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] *REDC* 419.

⁵⁹ Draft Data Act, explanatory memorandum [3] 13.

⁶⁰ David Harvey, *The New Imperialism* (OUP 2003).

for relegating accumulation based upon predation and violence to an ‘original stage’ that they considered outside of the capitalistic system – so-called primitive accumulation.⁶³ In Marxist terms, primitive accumulation is the prehistory of capital as it is the ‘*historical process of divorcing the producer from the means of production.*’⁶⁴ The capitalist system presupposes the ‘complete separation of the labourers from all property in the means by which they can realize their labour.’⁶⁵ To achieve such separation – in other words, to allow capitalists to own the means of production and subjugate labourers – requires considering the history of violent dispossession that is rooted in the enslavement of feudalism, colonialism, and the enclosures that created a landless proletariat.⁶⁶ This primitive accumulation, albeit important to understand capitalism, is not the result of the capitalistic mode of production; according to Marx, it is its starting point.⁶⁷ This is where Harvey differs, and we would concur. His phrase ‘accumulation by dispossession’ intends to underline the persistence of predatory practices in terms of accumulation of capital: it is a call for a ‘general re-evaluation of the continuous role and persistence of the predatory practices of ‘primitive’ or ‘original’ accumulation within the long historical geography of capital accumulation.’⁶⁸

At the core of digital dispossession is the attempt by 4IR companies to appropriate and otherwise control both the data that this system produces and the algorithms that underpin the system. Data and algorithms generated by users through ‘smart’ devices and through other interactions with 4IR networks and infrastructures are subject to dispossession in a variety of ways. IP laws equip these enterprises with a number of IP rights that can be

⁶¹ Karl Marx, *Il Capitale* (1867), vol 1 (Bruno Maffi tr, Aurelio Macchioro and Bruno Maffi, UTET 2008) ch 24.

⁶² Rosa Luxemburg, ‘The Accumulation of Capital: A Contribution to an Economic Explanation of Imperialism (1913)’ in Peter Hudius and Paul Le Blanc (eds), Nicholas Gray (tr), *The Complete Works of Rosa Luxemburg*, vol II: Economic Writings 2 (Verso 2016) 7.

⁶³ Harvey (n 59) 145. Cf Jim Glassman, ‘Primitive Accumulation, Accumulation by Dispossession, Accumulation by “Extra-Economic” Means’ (2006) 30 *Progress in Human Geography* 608.

⁶⁴ Marx (n 60) 898. Emphasis added.

⁶⁵ *ibid* 897.

⁶⁶ Unlike Marx, Coulthard argued that dispossession was not a singular event but a set of persistent and enduring practices of state violence that multiply coerced Indigenous peoples into the nation-state’s colonial project. See Glen Sean Coulthard, *Red Skin, White Masks: Rejecting the Colonial Politics of Recognition* (University of Minnesota Press 2014).

⁶⁷ Marx (n 60) 896.

⁶⁸ Harvey (n 59) 144. Such persistence was also noted by Michael Perelman, *The Invention of Capitalism: Classical Political Economy and the Secret History of Primitive Accumulation* (Duke University Press 2000).

used (and abused) in order to de facto own 4IR data, most notably database rights and trade secrets.⁶⁹ As these rights are perceived as not providing a sufficient layer of protection, 4IR companies resort to a complex web of technological protection measures, contracts, and factual control to put in place novel extractive practices that can negatively affect citizens, who are often unaware of them. One can think of the distress caused by finding out that Google was circumventing the ‘do not track’ signal, which was at issue in *Vidal-Hall v Google Inc.*⁷⁰ Or of the largest civilian surveillance network in US history created by Amazon’s Ring, the smart security system that was found to share a wealth of granular information with countless law enforcement agencies.⁷¹ While privacy harm is more immediately apparent, other and often more profound harms can come from digital dispossession. As a study of the neoliberal smart city showed, ‘data lies at the heart of most power relations today.’⁷² With this in mind, it becomes clear that dispossessing users of their data and creating enclosed 4IR systems – whose data is immune to access and re-use – lies at the heart of the principles of autonomy and self-determination. This can be seen with clarity in 4IR companies’ attempts to exploit users’ biases and vulnerabilities in order to manipulate them.⁷³ Digital dispossession goes to the core of our dignity, especially when 4IR data includes protected characteristics that allow companies to discriminate against certain categories of citizens. Going back to the example of Amazon’s Ring, a number of reports suggest that sharing video footage with the police is leading to a disproportionate number of incidents involving people of colour.⁷⁴

⁶⁹ These practices may be in violation of data protection and privacy laws, but that is beyond the scope of this chapter. See Guido Noto La Diega, ‘The Internet of Loos, the General Data Protection Regulation, and Digital Dispossession under Surveillance Capitalism’ in *Internet of Things and the Law: Legal Strategies for Consumer-Centric Smart Technologies* (Routledge 2022) ch 5.

⁷⁰ [2015] EWCA Civ 311.

⁷¹ Justine Morris, ‘Surveillance by Amazon: The Warrant Requirement, Tech Exceptionalism, & Ring Security’ (2021) 27 BUJ Sci. & Tech. L. 237.

⁷² Evgeny Morozov and Francesca Bria, ‘Rethinking the Smart City. Democritizing Urban Technology’ (2018) Rosa Luxemburg Stiftung New York Office 53 <<http://www.rosalux-nyc.org/rethinking-the-smart-city/>> accessed 17 November 2020.

⁷³ As the European Court of Human Rights held in *Satakunnan Markkinapörssi Oy and Satamedia Oy v Finland* (2018) 66 EHRR 8 [137], the European Convention of Human Rights (ECHR) ‘provides for the right to a form of informational self-determination’ (Art 8).

⁷⁴ This is not entirely new as surveillance has always disproportionately affected people of colour, as seen in Simone Browne, *Dark Matters: On the Surveillance of Blackness* (Duke University Press 2015).

While Marx may not be of help in order to understand the persistence of dispossession, it does come in handy in order to understand how the data-fuelled relationship between us and 4IR technologies is changing us in ways that are still impossible to fully grasp. In the *Economic Manuscript 1861-63*,⁷⁵ Marx explains how technology determined the passage from manufacture to industrial factory. In the former, the worker exercises their skill to control the tools used in the manufacturing process. In the latter, the industrial worker becomes the tool of a machine. The all-consuming labourer-machine relationship isolates the worker, who confronts ‘capital as an isolated individual, standing outside the social connection with (their) fellow workers.’⁷⁶ Isolated workers become ‘merely the living accessories (...) of the unconscious but uniformly operating machinery.’⁷⁷ Under 4IR capitalism, this isolation and passivity of workers is worsened by the fact that the machine is no longer only the ‘external body’ of the labourer when working in the factory: the machine is all around us, in our smart cities; it reaches our most private spaces, in the smart home; and enters our own body under the guise of smart health. This means that the isolation and passivity caused by dispossession adversely affects all of us, regardless of whether we qualify as workers. In other words, in the 4IR society, where data is the most sought-after commodity, citizens become round-the-clock workers inasmuch as they produce big data – thus generating value – whether they are aware of it or not.

4. Enter the sui generis right: a suitable incentive for sustainable data practices?

As seen above, in light of the rise of private ordering practices revolving around data-driven contracting, technological protection measures, and factual power, IP is not the only tool (ab)used by 4IR companies to maximize digital dispossession at the expense of a more sustainable future.⁷⁸ However, IP can contribute to these practices and, as such, it is useful to focus on a generally less well-known IP right: the sui generis right, also known as the

⁷⁵ Karl Marx, ‘Economic Manuscript of 1861-63. A Contribution to the Critique of Political Economy’ in Karl Marx and Friedrich Engels, *Collected Works*, vol 30 (Progress 1988).

⁷⁶ *ibid* 478.

⁷⁷ *ibid* 489.

⁷⁸ On the decreased role of traditional IP mechanisms to incentivize innovation in the algorithmic society see Marco Ricolfi, ‘Il Futuro Della Proprietà Intellettuale Nella Società Algoritmica’ [2019] *Giur it* 10.

database right. Often accused of creating unjustified monopolies in data and in collection of facts,⁷⁹ this right – and its exceptions – can be pivotal to opening up data in the 4IR.

The starting point is to understand whether 4IR data can be protected by the *sui generis* right. This comes with the proviso that, as 4IR is an umbrella term that encompasses a diverse range of technologies, it is not possible to say that in all circumstances or, conversely, in none, the relevant data will fall within the scope of this regime. However, while the prevalent view seems to be that no machine-generated data is covered by the *sui generis* right,⁸⁰ it can be argued that nothing in the law prevents machine data from being controlled through this right, and that at least in some scenarios its applicability can be regarded as instrumental to opening up the relevant data.⁸¹

The main innovation introduced by the Database Directive⁸² – the *sui generis* right – is the right to prevent extraction or re-utilization of the whole or a substantial part of the contents of a database.⁸³ The concept of ‘database’ is not limited to what one may naturally think of as a database (such as WestLaw) as it includes a variety of things ranging from booking websites⁸⁴ through job advert boards⁸⁵ to topographic maps.⁸⁶ A database is any (i) collection of independent works, data or other materials, that is (ii) arranged in a systematic or methodical way and is (iii) individually accessible by electronic or other means.⁸⁷ So, for example, the collection of voice interactions recorded by a Google Home device, stored by Google, which organizes it and makes it searchable, may qualify as a database.

4.1. *Creating v obtaining data*

⁷⁹ P Bernt Hugenholtz, ‘Abuse of Database Right. Sole-Source Information Banks under the EU Database Directive’ in François Lévêque and Howard Shelanski (eds), *Antitrust, patents and copyright: EU and US perspectives* (Edward Elgar 2005) 203.

⁸⁰ JIIP and Technopolis Group, ‘Study in Support of the Evaluation of Directive 96/9/EC on the Legal Protection of Databases’ (European Commission 2018) Final report.

⁸¹ Guido Noto La Diega, ‘Artificial Intelligence and Databases in the Age of Big Machine Data’ (2019) 25 AIDA 2018 93.

⁸² Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases [1996] OJ L 77/20.

⁸³ Database Directive, Art 7.

⁸⁴ Case *Ryanair Ltd v PR Aviation BV* [2015] All ER (EC) 662.

⁸⁵ Case C-762/19 *SIA “CV-Online Latvia” v SIA “Melons”* (CJEU, 3 June 2021).

⁸⁶ Case C-490/14 *Freistaat Bayern v Verlag Esterbauer GmbH* [2015] 10 WLUK 766.

⁸⁷ Database Directive, Art 1.

A database attracts sui generis protection if a substantial investment has been involved in obtaining, verifying or presenting its contents.⁸⁸ This requirement refers to an investment in terms of human, technical, or financial resources.⁸⁹ Regardless of its literal meaning, the threshold of substantiality is usually interpreted as inclusive of even low-level investment.⁹⁰ One could think of the technical investment in setting up a blockchain network or of the human investment in the cost of data labelling in supervised and semi-supervised machine-learning systems. While in most instances the substantiality of the investment is not in question, a major jurisprudential hurdle is constituted by the definition of ‘obtaining’ data. Indeed, according to the Court of Justice in *British Horseracing Board v William Hill Organization*⁹¹ and in the three *Fixtures Marketing*⁹² rulings of November 2004, only the investment in ‘obtaining’ existing data – construed as the opposite of ‘creating’ new data – counts towards sui generis protection. In other words, ‘investment in obtaining’ must be understood as referring 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 materials.’⁹³ On the face of it, this ‘obtaining-creating dichotomy’ would seem to exclude 4IR data from the scope of this regime. However, the dichotomy can be criticized, and arguably overcome, based on two considerations.

First, the dichotomy goes against the objectives of the Database Directive. The main goal of the Directive was to stimulate investment in the database sector, thus bridging the gap between the EU and the US database industries.⁹⁴ This fact must be analysed in light of empirical evidence clearly showing that the majority of investments by database owners concern data collection, rather than actually setting up the database.⁹⁵ The Directive itself refers to the importance of investment in both information storage and information

⁸⁸ Database Directive, Art 7(1).

⁸⁹ Case C-202/12 *Innoweb v Wegener* [2014] *JiPLP* 458.

⁹⁰ European Commission, ‘Staff Working Document “Evaluation of Directive 96/9/EC on the Legal Protection of Databases”’ (2018) SWD(2018)147 final 27.

⁹¹ Case C-203/02 *British Horseracing Board v William Hill Organization* [2004] ECR I-10415.

⁹² Case C-338/02 *Fixtures Marketing v Svenska Spel* [2004] ECR I-10497; case C-444/02 *Fixtures Marketing v OPAP* [2004] ECR I-10549; case C-46/02 *Fixtures Marketing v Oy Veikkaus Ab* [2004] ECR I-10365.

⁹³ *British Horseracing* (n 80) [31]; similarly, *Svenska Spel* (n 81) [39]; *OPAP* (n 81) [39-40]; *Oy Veikkaus* (n 81) [33-34], [41-42], [44-46], [49].

⁹⁴ Database Directive, Rec 11.

⁹⁵ European Commission (n 89) 36.

processing systems.⁹⁶ The express reference in Recital 10 to information that is ‘generated’ may be seen as encompassing both created and obtained data. Additionally, as the Directive aims at providing ‘a stable and uniform legal protection for databases,’⁹⁷ it could be argued that the sterilizing effect of the dichotomy – which brings outwith the scope of most databases – and the legal uncertainty that ensued do not achieve the envisaged stability of protection.

Second, and perhaps more importantly from this chapter’s perspective, the 4IR – with its fusion of cyber-physical and bio-technologies – calls into question the tenability of the obtaining-creating dichotomy. Even in a pre-smart world, the line between creating and obtaining data was blurred. This is perhaps best illustrated by the contested status of live football data, deemed to be ‘obtained’ in the UK,⁹⁸ and ‘created’ in Germany.⁹⁹ Even in the literature, there is no consensus on where to draw the line.¹⁰⁰ The dichotomy becomes increasingly meaningless in the 4IR age. One need only consider neural networks used to spot the correlation between existing data, spot patterns, and make predictions. Arguably, such inferences and relational data can be regarded as both created and obtained, both new and existing, which begs the question whether we had better abandon the distinction altogether. Another example could be data generated by the distributed sensors of a smart city (for instance, traffic sensors, pollution sensors). Usefully, in *Football Dataco v Sportradar* the Court of Appeal observed that ‘a scientist who takes a measurement would be astonished to be told that she was creating data’;¹⁰¹ in reality, the scientist is creating a record of data, not the data itself.¹⁰² The same could be said for smart city data, as well as for similar data produced by fish tags and nanostructured biosensors for smart agriculture.¹⁰³

4.2. *Exceptions and private ordering*

⁹⁶ Database Directive, Rec 12.

⁹⁷ *ibid.*

⁹⁸ *Football Dataco Ltd v Sportradar GmbH* [2013] EWCA Civ 27.

⁹⁹ European Commission (n 89) 25.

¹⁰⁰ See, eg, Estelle Derclaye, ‘Databases" Sui Generis" Right: Should We Adopt the Spin-off Theory?’ (2004) 26 EIPR; Andrea Ottolia, *Big Data e Innovazione Computazionale* (Giappichelli 2017) 79.

¹⁰¹ (n 87) [39].

¹⁰² *ibid* [57].

¹⁰³ cf Amina Antonacci and others, ‘Nanostructured (Bio)Sensors for Smart Agriculture’ (2018) 98 *TrAC Trends in Analytical Chemistry* 95.

Once having established that 4IR cannot be excluded from *sui generis* protection as a matter of principle, we can proceed to explain the benefit of this interpretive approach. Alongside rights for the database maker, a number of user freedoms can be leveraged to open up and share the database for the common good. Most notably,¹⁰⁴ the Copyright in the Digital Single Market (CDSM) Directive¹⁰⁵ has introduced text and data mining (TDM) exceptions, which refer to ‘any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations.’¹⁰⁶ The TDM exception could be used for more sustainable data governance, for example, to use AI and mine privately-held big data in order to detect patterns in climate change and possibly prevent some of its most extreme consequences. Positively, under Article 4 of the CDSM Directive, this is a mandatory¹⁰⁷ and open-ended¹⁰⁸ exception which requires Member States to allow reproductions and extractions of the contents of databases, and to retain them for as long as necessary for the purposes of TDM.¹⁰⁹ The main limitations

¹⁰⁴ Other exceptions and user freedoms may be of help. For example, the end-user of a database has the right to extract and re-utilize insubstantial parts of its contents for any purpose whatsoever (Database Directive, Art 8). This provision has the benefits of being qualified expressly as a user right and not as an exception (which makes its broader interpretation more likely), as well as being mandatory and binding. If AI can be programmed to extract data from several databases, with the proviso that the extraction will be ‘insubstantial’ with regard to each of them taken in isolation, then arguably one could rely on this right to re-use privately held data for the common good.

¹⁰⁵ Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (CDSM Directive) [2019] OJ L 130/92. Although the CDSM Directive was due to be transposed by June 2021, many Member States have not yet complied, hence the infringement procedure opened by the European Commission in July 2021. As of January 2022, Belgium, Bulgaria, Greece, Cyprus, Latvia, Luxembourg, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, and Sweden have not transposed this directive.

¹⁰⁶ CDSM Directive, Art 2(2).

¹⁰⁷ A significant innovation of the new directive is that Member States have an obligation to introduce the relevant exceptions in their domestic legal systems. Cf. Database Directive, Art 8: ‘*Member States may stipulate that...*’ (conversely, the user rights under Art 8 are to be regarded as mandatory).

¹⁰⁸ Cf. CDSM Directive, Art 3 (TDM for purposes of scientific research) and UK Copyright Designs and Patents Act 1988 (CDPA), s 28A (text and data analysis for the ‘sole purpose of research for a non-commercial purpose’).

¹⁰⁹ CDSM Directive, Art 4(1)(2). Compared to existing national exceptions, this passage can be interpreted as meaning that copies can also be transferred to other users (eg, other members of a team) as long as they are using the copies for TDM purposes. Cf. CDPA, s 29A(2), whereby copyright is infringed if a copy made for TDM purposes is transferred to any other person.

of this exception, however, are that it allows private ordering, namely use of contracts and technological measures to (over)protect IP. Indeed, unlike other exceptions, contractual provisions contrary to Article 4 are enforceable,¹¹⁰ and, even more worryingly, this exception does not apply if rightholders reserve the use of protected contents ‘in an appropriate manner, such as machine-readable means.’¹¹¹

Technological protection measures constitute the main obstacle to the exercise of user freedoms. Adopting the model provided by the US Digital Millennium Copyright Act, the Infosoc Directive incentivized rightholders, including database makers, to adopt effective technological measures (for instance, copy control mechanisms and other ‘locks’) to prevent unauthorized acts.¹¹² This form of ‘paracopyright’¹¹³ has been rightly criticized for leading to excessive monopolies as these technological protection measures can be used to prevent otherwise lawful uses, such as exceptions, limitations, and the public domain.¹¹⁴ In the EU, this risk is partly mitigated. As held in *Nintendo v PC Box*,¹¹⁵ technological protection measures should not go beyond prevention of unauthorized restricted acts¹¹⁶ and that, pursuant to the principle of proportionality, they should be designed so as to minimize interference with the rights and freedoms of third parties.¹¹⁷ This is a general principle that should also apply to the aforementioned TDM exception, which therefore should be interpreted as not allowing disproportionate technological protection measures. This rule becomes explicit with regard to the sister TDM exception that is limited to TDM carried out by research organizations and cultural heritage institutions for scientific research purposes under Article 3 CDSM. While we do not need to delve into this exception – as its subjective and objective scope makes it of little help in the context of general re-use of data for sustainability purposes¹¹⁸ – one need only mention that rightholders can only apply measures

¹¹⁰ CDSM Directive, Art 7(1).

¹¹¹ CDSM Directive, Art 4(3).

¹¹² Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society [2001] OJ L 167/10, Art 6(3).

¹¹³ Joseph P Liu, ‘Paracopyright – a Peculiar Right to Control Access’ in Jane C Ginsburg and Rochelle Cooper Dreyfuss (eds), *Intellectual Property at the Edge: The Contested Contours of IP* (CUP 2014) 227.

¹¹⁴ Lawrence Lessig, *Free Culture: The Nature and Future of Creativity* (Penguin 2004) 162.

¹¹⁵ Case C-355/12 *Nintendo Co Ltd v PC Box Srl* [2014] 1 WLUK 506.

¹¹⁶ *ibid* [31].

¹¹⁷ *ibid* [32].

¹¹⁸ The main limitations of this exception are that it can only be relied upon (i) by research organizations and cultural heritage institutions, and (ii) for scientific research purposes (CDSM Directive, Art 3(1)). See Alain

to ensure the security and integrity of the networks and databases where the contents are hosted, but ‘such measures shall *not go beyond what is necessary to achieve that objective*.’¹¹⁹ This can be regarded as an application of the general principle of minimum interference as expressed in *Nintendo v PC Box*. The principle is further specified in relation to the infamous upload filter. Under Article 17 CDSM, online content-sharing service providers can be found liable for copyright infringement if they do not put in place ‘best efforts to prevent (the) future uploads’¹²⁰ of allegedly infringing materials. While it is foreseen that platforms will enforce this provision by technological means – which raise more than one concern¹²¹ – these ‘best efforts’ cannot lead to ‘prevention of the availability of works or other subject matter uploaded by users, which do not infringe copyright and related rights, including where such works or other subject matter are covered by an exception or limitation.’¹²²

While we posit that the exceptions to the sui generis right can be invoked to open up 4IR data for sustainability purposes, this strategy has two shortcomings: one pertains to the unjustifiably narrow interpretation of the scope of application of the Database Directive in *Ryanair v PR Aviation*;¹²³ the other to exacerbation of the issue of IP overlaps in the 4IR.

4.3. *The problem with Ryanair v PR Aviation*

In this case, the defendant was a website operator allowing consumers to search through the flights of low-cost air companies, compare conditions, and book a flight. Its meta-search engine obtains the data in an automated way from a dataset linked to the Ryanair website.

Strowel and Rossana Ducato, ‘Artificial Intelligence and Text and Data Mining. A Copyright Carol’ in Eleonora Rosati (ed), *The Routledge Handbook of EU Copyright Law* (Routledge 2021) ch 15.

¹¹⁹ CDSM Directive, Art 3(3).

¹²⁰ CDSM Directive, Art 17(4)(c).

¹²¹ Christophe Geiger and Bernd Justin Jütte, ‘Platform Liability Under Art. 17 of the Copyright in the Digital Single Market Directive, Automated Filtering and Fundamental Rights: An Impossible Match’ (2021) 70 GRUR International 517; Sevrà Guler Guzel, ‘Directive on Copyright in the Digital Single Market and Freedom of Expression: The EU’s Online Dilemma’ in Tatiana-Eleni Synodinou and others (eds), *EU Internet Law in the Digital Single Market* (Springer International Publishing 2021).

¹²² CDSM Directive, Art 17(7). While a subsequent sentence refers to only some copyright exceptions (eg, quotations), this is a rule of general application as confirmed by the general obligation upon Member States to ‘ensure that users have access to a court or another relevant judicial authority to assert the use of an exception or limitation to copyright and related rights’ (CDSM Directive, Art 17(9)).

¹²³ (n 73).

The defendant's screen scraping, that is, the automated extraction of data from a website,¹²⁴ was in violation of the Terms & Conditions¹²⁵ that put in place an exclusive distribution system and prevent unauthorized websites from selling Ryanair flights. The use of the website was limited to private non-commercial purposes. The defendant's arguments, upheld by the referring court, revolved around the fact that national implementation of the Database Directive provided some limitations to contractual autonomy. In particular, the rightholder cannot prevent a lawful user from accessing the contents and making normal use of them¹²⁶ and this exception cannot be contractually overridden.¹²⁷ However, the referring court noted that the database was not protected because of lack of originality and of substantial investment. Therefore, the question to the Court of Justice was whether the scope of the Database Directive covered unprotected databases and, consequently, whether the limits on contractual freedom resulting from non-contractually-overridable exceptions and users' rights¹²⁸ also apply to such databases.¹²⁹

According to the Court, it is immaterial that Ryanair's database matches the definition of database given in the Directive. The latter provides for two different sets of rights and obligations that apply only if the relevant criteria are met (originality and substantial investment). Therefore, in the Court's reasoning, if the database maker does not have a right under the Directive, then exceptions cannot be invoked against them. The conclusion¹³⁰ is that the Database Directive is not applicable to a database which, despite matching the definition of a database, does not meet the requirements for copyright or sui generis protection. Therefore, under EU law there are no limits to the rightholder's freedom to lay

¹²⁴ cf Marco Caspers and others, 'Baseline Report of Policies and Barriers of TDM in Europe' (2016) Future TDM D3.3+ 9.

¹²⁵ The current version of the Terms & Conditions, as updated on 6 September 2021, no longer contains such provisions. See 'General terms and conditions of carriage' <<https://www.ryanair.com/gb/en/useful-info/help-centre/terms-and-conditions>> accessed 31 January 2022.

¹²⁶ Auteurswet, Art 24a(1), which corresponds to Article 6(1) of the Database Directive. The Court of Justice does not refer to the homologous provision applicable to the sui generis right, recognizing the lawful user's right to insubstantial extraction and re-utilization of the contents of the database (Database Directive, art 8).

¹²⁷ Auteurswet, Art 24a(3), corresponding to Database Directive, Art 15.

¹²⁸ The wording is slightly different even though the content is similar. On the one hand, there are exceptions to the database copyright (Art 6); on the other hand, the rights and obligations of a lawful user of databases are covered by the sui generis right (Art 8).

¹²⁹ *Ryanair* (n 73) [28].

¹³⁰ *Ryanair* (n 73) [49].

down contractual limitations on its use by third parties, because the exceptions under the directive will not apply.

The reasoning is open to criticism. Contrary to what the Court opined, Article 1 does not merely define the concept of ‘database.’ Rather, it expressly determines the *scope* of the Directive. Any database in any form will fall within the scope of the Directive, as long as the materials are independent, arranged systematically or methodically, and individually accessible. Therefore, since Ryanair’s website falls within the scope, one should recognize that the exceptions provided in the Directive prevail over the clauses in the website’s T&Cs. The difficulty of which provision to apply, whether the one on copyright or the one on the sui generis right, is only a practical one that could be overcome – for instance, finding the highest common denominator between Article 6 and Article 8, which does not seem complicated since the provisions are indeed quite similar. A different solution would, indeed, be unreasonable because it would lead to recognizing stronger protection to those databases where the database maker or author did not input any intellectual effort or any meaningful investment.

4.4. *Intellectual property overlaps*

The second limitation of an exception-focused strategy relates to the issue of IP overlaps, which becomes more pressing in the 4IR. Due to proliferation of IP rights, as well as to the widening of their subject matter and of the range of restricted acts, we have increasingly been witnessing the phenomenon of IP overlaps.¹³¹ This can take many forms, but is perhaps most clearly illustrated by the existence of two distinct IP rights covering the same product, for instance a computer program being protected under both copyright law and patent law.¹³² It is often difficult to deal with the interplay between different IP rights because the relevant laws tend to ignore the problem. For example, as noted in *Football Dataco v Sportradar*,¹³³ there is ‘nothing in the (Database) Directive which says that if a collection of data otherwise qualifying for a database right is contained within a copyright work, that collection cannot

¹³¹ Estelle Derclaye and Matthias Leistner, *Intellectual Property Overlaps: A European Perspective* (Hart 2011); Neil Wilkof and Shamnad Basheer (eds), *Overlapping Intellectual Property Rights* (OUP 2012).

¹³² Gustavo Ghidini and Emanuela Arezzo, ‘Dynamic Competition in Software Development: How Copyrights and Patents, and Their Overlapping, Impact on Derivative Innovation’ (2013) 3 Queen Mary Journal of Intellectual Property 278.

¹³³ (n 87).

also be the subject of a sui generis right.’¹³⁴ It is easy to see how the 4IR makes things worse. In the 4IR, we can observe the growing fusion between hardware, software, digital content, services, data, and biotechnologies. One need only think of how virtually every aspect of a smartphone will be protected by some form of IP, from the copyright on the code through the design rights on the shell to patents on some of the internal machinery, trade marks on the logos, and trade secrets on the algorithms.¹³⁵ This risks neutralizing the potential of IP exceptions because many of the acts covered by an IP right’s exceptions constitute infringement of another IP right. For example, whilst users can invoke a harmonized right to reverse engineer without the copyright holder’s permission, 4IR companies may block it by qualifying it as patent infringement.¹³⁶ Correspondingly, copyright holders’ power to control derivative software is at odds with the right to patents’ derivative non-obvious inventions. Finally, one needs to consider that IP exceptions tend to be misaligned. Some exceptions are mandatory, others are left to the discretion of Member States as to whether to implement them; some are binding, others can be overridden contractually; some cover unauthorized commercial uses, others do not; some are regarded as user rights, others are not. Whereas *de lege ferenda* this problem may be tackled – for example, through a treaty on IP exceptions¹³⁷ – as things stand the existence of overlapping IP rights that cover 4IR technologies, and the divergence between what acts are permitted without the owner’s consent hinders the potential of the TDM exception, and of IP exceptions more generally, to be used in order to open up data for the common good. Indeed, when overlaps occur – and in the 4IR they are set to

¹³⁴ *ibid* [29].

¹³⁵ See Guido Noto La Diega, ‘The Internet of Things (You Don’t Own) under Bourgeois Law. An Integrated Tactic to Rebalance Intellectual Property’, *The Internet of Things and the Law. Legal Strategies for Consumer-Centric Smart Technologies* (Routledge 2022) ch 6.

¹³⁶ In a decision that shows a lack of awareness of the excesses of private ordering, the Court of Justice recently held that the user of a computer program is only allowed to decompile it in compliance with any conditions set out in the end user licence agreement (case C-13/20 *Top System SA v Belgium* [2022] ECDR 4 [74]). Whilst software companies should not interpret this as carte blanche to ban decompilation contractually, it is likely that this ruling will be used to substantially decrease the availability of this right.

¹³⁷ This could build on the Instrument on Permitted Uses in Copyright Law drafted in January 2021 by the Max Planck Institute for Innovation and Competition <<https://www.ip.mpg.de/en/research/research-news/permitted-uses-in-copyright-law-research-team-develops-international-instrument.html>> accessed 1 February 2022.

become the rule rather than the exception – the strictest regime is likely to override the more generous one.¹³⁸

The convergence between IP overlaps and the rise of the 4IR acts as a powerful reminder that we should endeavour to understand IP – and the technologies which IP governs, or fails to govern – holistically. In practice, this means mainly two things. First, that IP scholarship needs to consider more closely the interplay between IP and the other data governance laws, meaning mainly the EU General Data Protection Regulation,¹³⁹ the Open Data Directive,¹⁴⁰ international trade law (for instance, the TRIPS plus provisions on data exclusivity),¹⁴¹ as well as the proposed Data Governance Act,¹⁴² and the AI Act.¹⁴³ Second, data generated by

¹³⁸ Estelle Derclaye, ‘Overlapping Rights’ in Rochelle Dreyfuss and Justine Pila (eds), *The Oxford Handbook of Intellectual Property Rights* (OUP 2018) 622.

¹³⁹ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (GDPR) [2016] OJ L 119/ 1. One of the most controversial aspects of the GDPR is to what extent IP considerations can be invoked to reject data subject access requests (Recital 63, Art 15(4)). See European Data Protection Board, ‘Guidelines 01/2022 on Data Subject Rights - Right of Access’ (2022) para 166; Noto La Diega and Sappa (n 57).

¹⁴⁰ Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information [2019] OJ L 172/56. Even though this directive does not apply to ‘documents for which third parties hold intellectual property rights’ (Art 1(2)(c)), research data has to be open by default and, when there are clashes with IP rights, it should be as ‘open as possible, closed as necessary’ (Art 10(1)).

¹⁴¹ ‘TRIPS plus’ refers to free trade agreements whereby the parties must introduce strong IP laws, beyond the minimum standard imposed by the Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPs). Considerable costs arise from implementing higher standards of protection (Carlos M Correa, ‘TRIPS and TRIPS-Plus Protection and Impacts in Latin America’ in Daniel Gervais (ed), *Intellectual Property, Trade and Development* (OUP 2014). This is particular evident when it comes to the data exclusivity provisions, which typically allow pharmaceutical companies to extend the duration of their monopolies beyond the expiration of the patent by keeping secret the trials data submitted for market approval. See, eg, US-Morocco Free Trade Agreement 2004, Art 15.10(2).

¹⁴² Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act or DGA) COM/2020/767 final. This proposal revolves around the fuzzy concept of ‘data altruism’ and it sets forth incentives and mechanisms for citizens and public entities to share their data in the general interest. The DGA also applies to ‘data which is not accessible due to commercial and statistical confidentiality and data for which third parties have intellectual property rights’ (Rec 7, see Art 3(1)(c)). Importantly, while re-use of data is said to be allowed only in compliance with IP rights, the DGA prevents public sector bodies from exercising their database rights ‘to prevent the re-use of data or to restrict re-use

and through 4IR systems does not exist in isolation. It would be pointless to focus solely on how to open up data if we do not also find ways to incentivize the other facets of the underlying technologies. Notably, open data strategies will have to be accompanied by efforts to realize free and open software, open hardware, open standards, and open platforms. While the law lags behind, bottom-up and community projects are already indicating this new path and fully embracing openness. Collective endeavours such as the Open Internet of Things Certification Mark, TrustableTech, DotEveryone, OpenTech UK, Arribada Restart Project, and the thousands of 4IR-focused meetups around the world, including in the Global South, tell a story of concerted efforts to open up the 4IR in order to make it more sustainable and societally beneficial. Interpretive strategies and law reform should learn from these efforts and identify adequate incentives to pursue openness and, ultimately, sustainability. This leads us to speak of the latest attempt to open data and deal with the interaction between the *sui generis* right and 4IR data, the proposed Data Act.

5. Policy options and the Data Act

On 23 February 2022, the European Commission unveiled its Proposal for a Regulation on harmonized rules and fair access to and use of data (Data Act). The draft establishes several rules that aim to ensure fairer access to and use of mainly non-personal data¹⁴⁴ (that is, data

beyond the limits set by this Regulation’ (Art 5(7)) Additionally, data covered by trade secrets can be re-used as long as no confidential information is disclosed as a result (Art 5(8)).

¹⁴³Under certain circumstances, market surveillance authorities can access data used for training, validation, and testing of AI systems, including the source code (Art 64). Additionally, high-risk AI systems have to be designed transparently ‘to enable users to interpret the system’s output and use it appropriately’ (Art 13(1)). The AI Act shows awareness that in some instances the imperative to access and re-use data will prevail over IP, with the proviso that interference has to be proportionate and limited to what is necessary to exercise the right to an effective remedy and to implementation of the duties of transparency towards supervision and enforcement authorities (Preamble, point 3.5).

¹⁴⁴ The current wording creates confusion as to what type of data is in scope. On the one hand, the Draft Data Act is set not to affect ‘the applicability of Union law on the protection of personal data’ (Art 1(3)); on the other hand, it applies to access to data defined broadly as ‘any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audio-visual recording’ (Art 2(1)). Importantly, the new law would apply to data holders, defined as ‘a legal or natural person who has the right or obligation, in accordance with this Regulation, applicable Union law or national legislation implementing Union law, *or in the case of non-personal data* and through control of the technical design of the product and related services, the ability, to make available certain data;’ which can be construed as implying that data holders are also in scope when processing personal data.

not protected under the GDPR and related instruments)¹⁴⁵ held mainly by big companies. The justification for the proposed Data Act is that if data were not underused it could be an ‘essential resource to secure the green and digital transitions,’¹⁴⁶ and that if the distribution of value from 4IR data were more balanced this would release ‘enormous potential *for boosting a sustainable data economy* in Europe.’¹⁴⁷ Against this backdrop, the proposed legislation has five objectives:

- 1) to facilitate access to and use of data by consumers and businesses including clarifying the Database Directive,¹⁴⁸
- 2) to enable use by public sector bodies of data held by big enterprises in exceptional circumstances,¹⁴⁹
- 3) to facilitate switching between cloud and edge services,¹⁵⁰
- 4) to put in place safeguards against unlawful data transfer without notification by cloud service providers,¹⁵¹ and
- 5) to mandate interoperability standards for data.¹⁵²

To these ends, the Data Act will introduce a right of users to access and use the data generated by use of their products and services,¹⁵³ a right to share it with third parties,¹⁵⁴ and an obligation for private companies to make data available to public sector bodies to satisfy

¹⁴⁵ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119/1; Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA [2016] OJ L 119/89; Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) [2002] OJ L 201/37.

¹⁴⁶ Draft Data Act, explanatory memorandum [1] 1; see also Rec 1.

¹⁴⁷ Draft Data Act, explanatory memorandum [1] 1.

¹⁴⁸ Draft Data Act, Chs 1–4 and 10.

¹⁴⁹ Draft Data Act, Ch 5).

¹⁵⁰ Draft Data Act, Ch 6.

¹⁵¹ Draft Data Act, Ch 7.

¹⁵² Draft Data Act, Ch 8.

¹⁵³ Draft Data Act, Art 4.

¹⁵⁴ Draft Data Act, Art 5.

exceptional needs,¹⁵⁵ including prevention of a public emergency,¹⁵⁶ which arguably should encompass prevention of climate catastrophe. In particular, the user right to share data and the obligation on companies to make data available to public sector organizations may play a crucial role in allowing re-use of 4IR data for sustainability purposes.

The *sui generis* right stands out as the only notable exception to the approach of the Data Act whereby new rights and obligations should not be allowed to trump the legal instruments provided in related laws, namely IPRs, trade secrets, and data protection.¹⁵⁷

The amended proposal has only one article affecting the *sui generis* right, namely Article 35 about ‘Databases containing certain data.’ It reads:

The *sui generis* right provided for in Article 7 of Directive 96/9/EC shall not apply when data is obtained from or generated by a product or a related service falling within the scope of this Regulation, in particular in relation to Articles 4 and 5.

The provision is accompanied by a single relevant recital (Recital 84) whereby:

In order to eliminate the risk that holders of data in databases obtained or generated by means of physical components, such as sensors, of a connected product and a related service claim the *sui generis* right under Article 7 of Directive 96/9/EC, and in so doing hinder in particular the effective exercise of the right of users to access and use data and the right to share data with third parties under this Regulation, it should be clarified that the *sui generis* right does not apply to such databases. That does not affect the possible application of the *sui generis* right under Article 7 of Directive 96/9/EC to databases containing data falling outside the scope of this Regulation provided the requirements for protection in accordance with Article 7(1) of that Directive are fulfilled.

The article seeks to prevent the Database Directive’s *sui generis* right from applying to databases made from data generated by a product or related service covered by the Data Act. A connected product is defined in Article 2 as:

an item, that obtains, generates or collects, data concerning its use or environment, and that is able to communicate product via a publicly available electronic communications service, a physical, connection or on-device access and whose primary function is not the storing, processing or transmission of data on behalf of third parties, other than the user

¹⁵⁵ Draft Data Act, Art 14.

¹⁵⁶ Draft Data Act, Art 15.

¹⁵⁷ In this sense see Josef Drexler and others, ‘Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission’s Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act)’ 5 <<https://www.ip.mpg.de/en/research/research-news/position-statement-on-the-eu-data-act.html>> accessed 7 November 2022.

and a related service as:

a digital service other than an electronic communications service, including software, which is connected with the product at the time of the purchase in such a way that its absence would prevent the product from performing one or more of its functions, or which is subsequently connected to the product by the manufacturer or a third party to add to, update or adapt the functions of the product

The Data Act seeks to clarify that the term ‘obtaining’ in Article 7 of the Database Directive does not encompass data generated through the use of connected products or related services so that the *sui generis* right does not arise, thus ensuring these databases fall outside the scope of the *sui generis* right’s protection. This is instrumental to the obligation for the data holder to make data available,¹⁵⁸ as well as the user’s right to access and use data generated by the use of products and related services and share it with third parties (Articles 4 and 5 Data Act). These rights and obligations will remain intact in relation to such machine-generated data (which we have also called 4IR data in this article), that is, unaffected by the *sui generis* right. Under the previous version of the data act, a combined reading of the recital and article was somewhat confusing. Under one reading, if the data could be said to be obtained by such connected products, the *sui generis* right could still apply if the conditions are fulfilled. This is because Recital 84 said ‘where such databases do not qualify for the *sui generis* right.’ The use of ‘where’ was confusing as it can be construed as a fact – databases made of data obtained by connected products are not protected by the *sui generis* right – or a normative statement – databases made of data obtained by connected products should not be protected by the *sui generis* right. This wording has been removed under the new version and the recital is much clearer.

The recital leaves intact the application of the *sui generis* right for databases containing data falling outside the scope of this Regulation. Indeed, Recital 14a of the Data Act explains that the act applies to raw data. Inferred or derived data is outwith the scope of the Data Act, which means that the *sui generis* right could cover 4IR databases containing inferential and derivative data. This interpretation would not be unprecedented¹⁵⁹ and might be justified by a purposive construction of the Act, whose aim is to ensure distribution of

¹⁵⁸ Draft Data Act, Art 3.

¹⁵⁹ For example, the right to data portability does not apply to inferred data and derived data according to Art 29 Working Party, ‘Guidelines on the Right to Data Portability’ (2017) WP242 rev.01 9.

value between the different players of the data economy.¹⁶⁰ However, it is open to a threefold objection.

Firstly, the justification for considering ‘generated, obtained, or collected data’ as in scope is that (1) this data represents the digitization of user actions and events; (2) it is valuable to the user; (3) it can support innovation and the development of ‘digital and other services *protecting the environment, health and the circular economy*.’¹⁶¹ Arguably, the same rationale may justify the inclusion of ‘derived and inferred data’ in the scope of the Act.

Secondly, the 4IR disrupts traditional boundaries (such as between cyber and physical, between good and service)¹⁶² and does not cope well with binary legal fictions, as we have seen with regard to the obtaining-creating dichotomy. Therefore, to replace one dichotomy with another does not seem the most sensible way forward.

Thirdly, Recital 14a excludes ‘information derived from this data, which is the outcome of additional investments into assigning values or insights from the data, in particular, by means of proprietary, complex algorithms, including those that are a part of proprietary software,’¹⁶³ from the scope of the Regulation. This, more clearly than the previous Recital 17, points in the direction of an exclusion of inferential data from the new rights and obligations, and its inclusion in the scope of the Database Directive.

Another unclear aspect in the initial Regulation proposal was the status of 4IR databases where the data is neither obtained nor generated but is only verified or presented by means of connected products or related services. As data will first be obtained or generated and only then verified and presented, this may not instantiate a frequent occurrence. The reworded Article 35 and Recital 84 make clear that such databases would also be falling outside the scope of the act if they attract the database *sui generis* right.

¹⁶⁰ The Act aims to ensure ‘fairness in the allocation of value from data among actors in the data economy and to foster access to and use of data’ (*explanatory memorandum* [1] 2).

¹⁶¹ Draft Data Act, Rec 14a.

¹⁶² See, eg, Guido Noto La Diega, ‘Software Patents and the Internet of Things in Europe, the United States and India’ (2017) 39 EIPR 173.

¹⁶³ Data Act, Rec 17. Due to its unclear wording, this recital may be interpreted as meaning that inferred data is excluded from the scope of the Data Act only inasmuch as its access and re-use would lead to a breach of IPRs over the software generating it. However, this interpretation would go against the passage in Rec 14 whereby derived/inferred data ‘should not be considered within scope of this Regulation.’

Another point illustrating an utter lack of clarity is the regime that applies to mixed databases; that is, those databases which include both data falling within the Data Act's scope and data which does not (such as derived/inferred data). Does the Data Act apply only to those parts of the database that fulfil its conditions? How does a user know this? Should the data holder 'tag' the elements subject to the Data Act and those subject to the sui generis right?

A further issue is that the concept of 'machine' and '4IR' is broader than the definition of 'product' under Article 2 of the Data Act. This means that machine-generated data that goes beyond the definition remains covered by the sui generis right if it fulfils the conditions set out in the Database Directive.

Finally, and crucially, to exclude (some or all) 4IR data from sui generis protection risks being useless and even dangerous if the EU lawmaker does not clearly deal with the issue of the role of contracts and technological measures. The risk, as epitomized by *Ryanair v PR Aviation*, is that if left without a clear legal framework, data is appropriated by contractual and technological means to the detriment of users' rights and freedoms. Unfortunately, Article 35 fails to state that the provision is imperative, which may be construed as meaning that technological protection measures and contracts may prevail. Some of the provisions of the Data Act positively set forth that the Act's rights and obligations prevail over technological protection measures (Article 11(1)¹⁶⁴ and contracts (Article 12(2)). However, the focus of Article 11(1) is on third-party rights, whereas it should be made incontestable that it applies to all user rights. The main limitation of Article 12(2) is that it only refers to data sharing agreements and to their use to exclude application of Chapter III of the Act ('obligations for data holders legally obliged to make data available'), while it remains silent as to other obligations and rights. It is also unclear how these interact with the sui generis right. This should be addressed in the final version of the Data Act.

In sum, the initial draft of the Data Act left the tension between data openness and IP largely unresolved.¹⁶⁵ When we originally drafted this chapter, we hoped that the Data Act would evolve to contain at least a better phrased Article 35 and Recital 84 and even more provisions to revise problematic aspects of the Database Directive addressing the issues

¹⁶⁴ Draft Data Act, Art 11(1).

¹⁶⁵ To the point that it has been argued that the Data Act may be construed as creating something akin to data ownership; see Simon Geiregat, 'The Data Act: Start of a New Era for Data Ownership?' (8 September 2022) <<https://papers.ssrn.com/abstract=4214704>> accessed 7 November 2022.

highlighted in our chapter and elsewhere.¹⁶⁶ Some of our pleas were heard but not all. In the main, the issue of the non-imperative character of Article 35 remains. The argument may still be put forward that the Database Directive itself should be revised separately rather than being revised within the Data Act. We do not think such an argument is insurmountable as the proximity in scope of the two instruments is so close and there is no written rule of EU law that says a directive cannot be revised within a regulation. However, it appears from conversations with DG Connect staff that the Commission does not intend to revise the Database Directive via the Data Act. It remains open to revising the Directive in future but when and to what extent is as yet unclear.

6. Conclusion: findings, recommendations, and suggestions for further research

Data and sustainability have become an inseparable pair. After years of conceiving data exclusively as a private matter – individual privacy and commercial interest to data monetization – the simultaneous rise of 4IR and of sustainability is leading to the requalification of data as an asset for the common good. In this chapter we have presented a threefold meaning of ‘data sustainability’ that accounts for (i) the use of data to measure progress towards the SDGs; (ii) the use of data-driven 4IR solutions to pursue sustainability targets; (iii) how current data practices hinder or foster the common good. Against this backdrop, we have denounced how 4IR companies tend to digitally dispossess end users and create data enclosures preventing citizens, other companies, and public sector organizations from accessing and re-using that data in the public interest. IP law could provide some solutions but, as our analysis of the *sui generis* right has demonstrated, it lacks clarity as to its applicability to 4IR databases and as to the scope of the relevant exceptions, which leads to increased digital dispossession by contractual and technological means. On a positive note, the proposed Data Act has signalled that better access and sharing of data could be in sight, but its reform of the *sui generis* right is far from providing much-needed clarity. If one of the Act’s goals was to overcome the ‘uncertainty (...) as to the accidental or unintended application of the *sui generis* right to databases containing machine-generated data,’¹⁶⁷ then it is fair to say that the goal has not been fully achieved. More balanced data governance and better access to and re-use of data is fundamental for private and public organizations to efficiently re-use data to contribute to the vital fight for climate justice and sustainable

¹⁶⁶ Estelle Derclaye and Martin Husovec, ‘Sui Generis Database Protection 2.0: Judicial and Legislative Reforms’ [2022] 6 EIPR 325-333.

¹⁶⁷ Draft Data Act, explanatory memorandum [3] 9.

innovation. This cannot be achieved if we allow 4IR to effectively claim ownership over valuable inferential data.

In the future, IP and information law studies should further explore the concept of ‘data sustainability;’ the latter cannot be reduced to an issue of environmental law or data governance. Future research should also holistically understand the relation between existing and new ‘access rights’, broadly construed, that seem to be proliferating (from the TDM exceptions¹⁶⁸ through the right to access the source code of high-risk AI systems,¹⁶⁹ to the power to access non-personal data¹⁷⁰). At the same time, the interactions between those rights, IP rights especially trade secrets, contracts, and technological measures still remain to be worked out.

¹⁶⁸ CDSM Directive, Arts 3–4.

¹⁶⁹ Draft AI Act, Art 64(2).

¹⁷⁰ Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union (‘Free-Flow of Non-Personal Data Regulation’) [2018] OJ L 303/59, Art 5.