



Tommaso Fia\*

# An Alternative to Data Ownership: Managing Access to Non-Personal Data through the Commons

<https://doi.org/10.1515/gj-2020-0034>

Published online September 23, 2020

**Abstract:** In today's algorithmic society, access to large-scale datasets is the *sine qua non* for any economic actor to reap the benefits of data-driven innovation (DDI). This article explores alternative mechanisms of data management in large-scale processing environments which can bolster access in view of the shortcomings of the existing data ownership-centric system. The scope of the analysis is limited to non-personal data. First, this contribution elaborates on the features and shortcomings of the data ownership-centric system and the existing legislation on data access. In fact, despite its ground-breaking potential, data access is not a widely available resource. It is subject, meanwhile, to the ability of several actors to control it, originating from data holders' position of *de facto* control over data ("data ownership"), which is mostly anchored in technological, behavioural, and legal access barriers. This ownership-oriented setting thus stifles data sharing and opportunities for novel reuses of data. Despite these concerns, EU secondary legislation and case law (including the "essential facilities doctrine" of competition law) have not yet offered appropriate means to enable data access across society. Second, this article investigates whether alternative systems of data management based on the commons is a viable solution to open up access to raw non-personal data (RNPd). The commons as a conceptual notion and institutional mechanism values access and freedom to operate, instead of power to appropriate. The article homes in on two main reasons which substantiate why commons management

---

The first half of the paper title pays tribute to the landmark book of a renowned Italian scholar (Grossi 1981). In this study, Grossi focuses on the debates which developed in the second half of the nineteenth century amongst legal scholars and social scientists over the roots and natures of property rights. This leads him to examine the forms of collective ownership which had evolved over the preceding centuries as a counterweight to the conception of absolute individual property, ensuing from a purely man-made and non-natural process. Obviously, this paper does not aim to match Grossi's historical and comparative analysis. However, it shares with the latter the intention of dwelling on communal institutional mechanisms other than private property, such as the commons. Hence the borrowing.

---

\*Corresponding author: Tommaso Fia, Ph.D. Researcher in Law, European University Institute, Firenze, Italy, E-mail: [tommaso.fia@eui.eu](mailto:tommaso.fia@eui.eu)

of RNPDP can be desirable. On the one hand, RNPDP can be deemed a cooperative infrastructural resource that calls for being pulled out of its factual enclosure (“structuralist approach” of the commons). On the other hand, grasping RNPDP as a commons means valuing its functional nature, making data available to a wide number of actors for the fulfilment of fundamental rights and enhancing human flourishing (“functionalist approach”). The article concludes with some thoughts on the lines of research which are still to be explored to put the commons-based vision of data management into practice.

**Keywords:** data ownership, the commons, data access, intellectual property law, EU law

## 1 Introduction

Over the last decade, new and enhanced information-based technologies have fuelled the digital economy. Today, the amount of data generated is far greater than it ever has been, for the cost of storing and processing data has significantly plummeted. Some figures are instructive in this respect. By 2025, the volume of data generated globally will reach 175 zettabytes (i.e. 175 trillion gigabytes), the rough equivalent of 67 million times the information enclosed in the collection of the Library of Congress of the United States (Commission 2020a, 2; OECD 2015, 20). Data feeds powerful algorithms, which turn it into valuable information and knowledge. This new phase, known as the data economy, means that companies and public bodies can orient their activities in light of data analytics and produce goods and services at (nearly) zero marginal cost (Rifkin 2014). Both these components form the “data driven innovation” (DDI) (OECD 2015, 21).

DDI is a new and cross-sectoral source of growth which revolves around collection, analysis, and reuse of incalculable amounts of data, which by itself is often of low value.<sup>1</sup> Big Data,<sup>2</sup> the Internet of Things (IoT),<sup>3</sup> Artificial Intelligence

---

<sup>1</sup> For the purposes of this paper, “data” means any digital representation (in the form of binary codes) which amounts to a lack of uniformity in the real world, whereas “information” indicates data having a particular meaning (semantic content) (Floridi 2010).

<sup>2</sup> Big Data technologies rest on (at least) four different features (the “4 Vs”): Velocity, Variety, Volume, and Veracity (De Mauro, Greco, and Grimaldi 2016).

<sup>3</sup> ‘The Internet of Things (IoT) creates an intelligent, invisible network fabric that can be sensed, controlled and programmed, in ways that enable artefacts to communicate, directly or indirectly, with each other and the Internet’ (Pagallo, Durante and Monteleone 2017, 60).

(AI)<sup>4</sup> and Machine Learning,<sup>5</sup> being the most prominent applications of DDI, are responsible for an extraordinary flourishing of many sectors of the economy. Numerous studies have pointed out and elaborated on the diverse applications, risks, and challenges of these data-exploitation technologies.<sup>6</sup> Notably, it is striking how the latter give the power, amongst other things, to optimise processes, boost productivity, and predict events, paving the way for simple identification of emergent societal needs, as well as market trends. Providing some topical examples in this respect can help understand their impact. Analysis of data hoarded by smart industrial machinery provides companies with valuable insights into the production of goods. Aggregation of consumers' information helps online businesses target populations of prospective customers. Sensor data harvested by city buses and trams is reused to streamline public transport lines according to users' needs and traffic flows. All these instances of everyday reality lead scholars to refer to this backdrop as a "revolution" (Floridi 2014; Mayer-Schönberger and Cukier 2013).

Despite its ground-breaking potential, data is no nearer to being a widely available resource. It is indeed subject to the *de facto* ability of several actors to control it and prevent others from doing so. This tendency, usually labelled "data ownership", is mostly based on technological measures which strengthen the position of data holders to the detriment of others that cannot access it.<sup>7</sup> Data ownership has far-reaching repercussions on society. When datasets slip into the hands of a few players (i.e. big techs, such as Google, Amazon, Facebook, Apple), power concentration in private actors has indeed quite no parallel in history. Not only does this predominance depend on economic power, but it also results in socio-political power (Ricolfi 2017, 218), which is the capacity 'to alter the behaviours, beliefs, outcomes, or configurations of some other entity' (Benkler 2016a, 19).

Against this backdrop, legal scholarship has investigated two lines of research concerning data ownership. Part of the literature has assessed

---

<sup>4</sup> It is rather difficult to rely upon a consensus definition of AI (OECD 2019a). According to Sartor, AI is the field aiming to develop computational methods of intelligent behaviours, so that computers and machines can perform tasks which usually require human intelligence to be carried out (Sartor 1996).

<sup>5</sup> "Machine [L]earning" refers to a subfield of computer science concerned with computer programs that are able to learn from experience and thus improve their performance over time' (Surden 2014, 89).

<sup>6</sup> The first systematic work on the impact of data-exploitation technologies on society is (Mayer-Schönberger and Cukier 2013). For a comprehensive analysis, see also (OECD 2015).

<sup>7</sup> On the notion of *de facto* ownership of personal and non-personal data, see, amongst others, Purtova (2015, 99–100), Ricolfi (2017), Ullrich (2019, 26–28), Ricolfi (n.d.).

whether existing intellectual property rights can protect data (Drexl 2017; Gervais 2019; Hugenholtz 2018; Wiebe 2017), and whether the creation of a new property right on data can facilitate data sharing across the markets.<sup>8</sup> Many legal scholars (Drexl 2017; Drexl et al. 2016; Hugenholtz 2018; Kerber 2016a, 2017, 127; Ricolfi n.d.) and the European Commission,<sup>9</sup> meanwhile, have stressed the importance of adopting measures aiming to enhance access to data. Most of these analyses, however, refrain from providing robust legal constructs on which access ought to be founded (Hugenholtz 2018; Kerber 2016b; Wiebe 2017). For instance, protection of initial investments of data holders in data production causes some literature to elaborate cautious approaches when it comes to granting data access to third parties (Kerber 2017, 119–20). Other jurists and the Commission interrogate how instruments of some bodies of law, such as competition law<sup>10</sup> or contract law (Janal 2017), can help to achieve broader data access. In doing so, the literature exploring data access generally tends to treat access as an exception, rather than a rule itself, and shies away from proposing effective adjustments of the law vis-à-vis the features of the data economy.

This contribution takes a first step in understanding whether different mechanisms of data management can bolster access in view of the shortcomings of the ownership-centric system. Yet, this paper does not aim to provide all-encompassing solutions which apply to any data processing practice. Application scenarios of data-exploitation technologies are too broad to be examined within a contribution as restricted as this one. Hence, two limitations circumscribe the analysis. First, it is confined to access to non-personal data in large-scale processing activities. Data protection and privacy obligations notably apply only to personal data.<sup>11</sup> They represent a legal barrier to its access (Kerber 2017) and stand as a bulwark of data subjects' interests by constituting a fundamental right. This aspect clearly requires a more challenging and articulate balancing of interests and rights which falls outside the scrutiny. Second, the paper leaves aside cases in which drawing a line

---

**8** This is the solution proposed by (Zech 2016). *Contra*, amongst others, Drexl et al. (2016), Kerber (2016a), Drexl (2017), Hugenholtz (2018).

**9** See, in particular, the documents issued by the EU Commission aiming to build the European data economy (Commission 2017) and create a “European data space” (Commission 2018).

**10** Issues pertaining to competition law are analysed by Commission (2017, 21–22). On the need for rejigging antitrust in light of the data economy, see also Ricolfi (2017).

**11** 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 L119/1, art 3(1). Hereafter ‘GDPR’.

between personal and non-personal data proves difficult.<sup>12</sup> The core<sup>13</sup> of non-personal data comprises different kinds of digital material: data generated by industrial machinery and farm equipment, sensor data produced in public transport, (real-time) data from online-accessible databases, and so forth. This contribution will deal only with the core and set aside the blurry penumbra of mixed datasets and other controversial cases.

The paper proceeds in four steps. Sections 2 through 4 analyse the shortcomings of the existing ownership-centric system of data management. These issues arise from (i) *de facto* data ownership, and (ii) a lack of legal rules enhancing access to non-personal data in large-scale processing environments. Each of the matters is investigated in Section 2, and Sections 3 and 4 respectively. Section 2 focuses on the sources of *de facto* ownership, i.e. access barriers. Section 3 draws a descriptive taxonomy of the data access regimes under the most salient EU legislation, bringing to light their inadequacies in enhancing data access. Section 4 centres on the essential facilities doctrine of competition law, which has been seen by part of the literature as a suitable framework for making data more available across the markets. The analysis shows, however, that this tool is unfit for the intended purpose. Section 5 then paves the way towards access-oriented modes of data management. It explores the application of the theories of the commons to raw non-personal data. Particular attention is paid to assess whether the commons is a feasible conceptual framework to enhance data access, and how it can inspire the making of a right to data access. Section 6 concludes.

## 2 *De Facto* Data Ownership: Features and Shortcomings

Currently, actors of the data economy rely on a management system of non-personal data based on *de facto* ownership. This system is anchored in the ability of appropriating large-scale datasets by restricting access to these resources. Access restrictions arise from a set of barriers affecting the links of the data value chain, i.e. collection, storage, analysis, and (re-)use (Rubinfeld and Gal 2017, 349ff). There exist three phenotypes of barrier: technological, behavioural, and legal.<sup>14</sup> Despite the particular traits of each typology, access barriers are closely interlinked. Each of them mutually reinforces and is corroborated by the others.

---

<sup>12</sup> The Commission has clarified that it can be rather difficult to draw such a distinction in respect to “mixed datasets”, in which personal and non-personal data are “inextricably linked”. In these cases, data protection law applies to all data forming the dataset (Commission 2019, 6–7).

<sup>13</sup> Reference here goes to the distinction between the “core” and the “penumbra” drawn by Hart (1958).

<sup>14</sup> This categorization is introduced by Rubinfeld and Gal (2017, 350).

Technological barriers pertain both to (i) relying on more advanced and efficient technologies than other players, and (ii) implementing technical protection measures (“TPM”). The first kind results in entry barriers which characterise data markets and are the object of study of microeconomics (e.g. economies of scale, economies of speed, multisided markets, positive network externalities, and so forth) (Rubinfeld and Gal 2017, 349ff). TPM typically consist in access-management technological tools granting protection of creative works in digital environments (“digital rights management”). However, since mass unstructured datasets are rather unfit for copyright protection (Farkas 2017, 8–9; Wiebe 2017, 64), related TPM do not enjoy legal safeguards.<sup>15</sup> In data environments, they equal to self-enforced limitation of data access through a merely technological enclosure (Mezzanotte 2017, 168; Ullrich 2019, 27).

Behavioural barriers arise in the presence of contractual limitations. In fact, most actors have recently switched from one-off contractual agreements to long-term relationships in which the provision of services comes with real-time data streams. In doing so, data providers and manufacturers grant licences that stop licensees from certain utilisation of non-personal data. This is particularly the case for exclusivity clauses which prevent purchasers of IoT devices (e.g. smart tractors or cars) from sharing data with third parties, re-using it for purposes other than the agreed ones (Tusikov 2019, 127), or even porting data once they intend to switch to another provider or manufacturer (Ricolfi 2017, 223–24). Exclusionary practices of this type are mostly anchored in an unequal bargaining power between the contractual parties (Drexel et al. 2017, 10). Other barriers are found in the terms and conditions (“Ts&Cs”) of websites. Ts&Cs generally restrict third parties from using bots and similar automated systems to hoard data from databases which can be accessed online (so-called “screen scraping”) (Surblytė 2016, 19–25).

Legal barriers to access stem from the allocative option of rights. IP and property-like rights on data epitomise this kind of access obstacles. Most of the literature agrees that only the quasi-IP protection of trade secrets fits massive unstructured sets of non-personal data (Zech 2016, 62–65; Ricolfi 2017, n.d.). As a private law scholar would argue, a trade secret is like the possession of a material object (i.e. factual disposal of a resource). Secret holders are likewise protected insofar as they take reasonable steps to keep information secret:<sup>16</sup> ‘[o]nce secrecy is lost, legal protection is lost as well’ (Wiebe 2017, 65) In this sense, holders of large-scale datasets typically maintain secrecy by

---

<sup>15</sup> TPM are legally recognised if targeted to prevent copyright-infringing behaviours (Directive 2001/29/EC on the harmonization of certain aspects of copyright and copyright related rights in the information society of 22 May 2001 [2001] OJ L 167/10, art 6(3)). See Ullrich (2019, 20–21).

<sup>16</sup> Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ L157/1, art 2(1)(c).

implementing organisational measures (e.g. access permissions to business premises), contractual arrangements (for example, non-disclosure agreements) and technical measures (for instance, data encryption) (Wiebe and Schur 2019, 818–21).

Against this backdrop, *de facto* data ownership brings about at least two shortcomings which can be easily imagined. One has already been mentioned in the foregoing and is about unequal power. In fact, smaller actors (such as SMEs and start-ups) cannot overcome imbalances in negotiating power (Commission 2020a, 8). An instructive example in this regard is the lock-in effect stemming from the exclusivity agreements between an IoT manufacturer and a purchaser. The other issue pertains to innovation. Again, competition analysis can lend a hand in this respect. Some authors point out that a management framework based on large-scale data exploitation is prone to winner-takes-all scenarios.<sup>17</sup> In short, data holders (incumbents) simply seize on DDI for their own purposes. Without any duty or incentive to share innovation, they keep it all for themselves. Most actors therefore typically generate and analyse data in-house. Even when they contract out data analytics practices, they do not make use of datasets for purposes other than the ones for which data were collected in the first place (Commission 2017, 9; Lohsse, Schulze, and Staudenmayer 2017, 15).

Within this framework, one may think that some pieces of legislation may already provide tools for surmounting these barriers. Several laws indeed allow (or impose, in some cases) access to non-personal data held by another entity. However, they have little relevance as a way of boosting data availability in the data economy. This is the other side of the coin – usually neglected by the literature – of the ownership-centric management system of non-personal data, which is addressed in Section 3.

### 3 Data Access and the Law: an Access Seeker-based Taxonomy

As signalled in the Introduction, numerous authors have frequently examined current trends of control over data. Nonetheless, a systematic picture of the legal regimes that could grant access to large-scale datasets is missing. This Section will shed some light in this respect by delineating a taxonomy. The latter illustrates how access mechanisms work under the most salient EU secondary legislation and underscores their gaps and shortcomings in enabling data access in the data economy.

To begin with, it is worth elaborating on what is meant by “access”. Analysing the legal sources of data access shows that there is no single definition. Empirical evidence

---

<sup>17</sup> This point is particularly stressed by Ricolfi (n.d.), who cites the far-seeing analysis of Spencer (2001).

may bring some clarity instead. The notion of access typically denotes a system of relations amongst people which qualifies “all possible means by which a person is able to benefit from things” (Ribot and Peluso 2003, 156).<sup>18</sup> It pertains to the opportunity to take advantage of things (being “things” both tangible and intangible commodities) (Ribot and Peluso 2003, 173), and consists in a “bundle of powers” over assets and resources (Crétois 2015, 323; Ribot and Peluso 2003, 158) whose allocation is commended to the law. However, rules of access to corporeal goods and immaterial objects are not alike. Whereas a homogenous body of law allots and governs access to the first category, i.e. property law, no clear-cut conclusion can be inferred when it comes to data. In fact, data access is mostly rooted in a plethora of sector-specific legislation. Although these access regimes cover different scopes of application, they can be scrutinised through the lens of their similarities as to shaping data access.

So, if access is in the spotlight, it is worth centring on *who* can access data. Accordingly, focusing attention on how laws govern entities seeking access (“access seekers”) from data holders (“access granters”) may be a way of drawing a sound taxonomy.<sup>19</sup> The departure point would thus be the organisational trait of those requesting access to data, taking into consideration the classical distinction between private and public sector. We then need to consider how open data access is constructed by the law (the “degree of access”), which depends on how many actors are enabled to obtain access. As a final step, we need to interrogate how rules practically design and shape data access from both a legal and technological stance (the “access approach”).<sup>20</sup>

### 3.1 Access Domains

Most legal regimes governing access to data revolves around a private-public sector differentiation which rests upon data controllership. The private or public nature of the entity holding and processing data generally defines the legislation’s scope of application. Hence, all data ‘that is generated, created, collected, processed, preserved, maintained, disseminated or funded by or for ... private corporations, households and non-profit institutions serving households private-sector data’ (OECD 2019b, 27) forms private-sector data. Data, information<sup>21</sup> and documents produced and held by public authorities are, on the other hand, included in public-sector data. The logic of this controllership-based classification can help us to discern access seekers, instead of data holders. Accordingly, private

<sup>18</sup> Ribot and Peluso refer to tangible goods.

<sup>19</sup> See examples in Table 1 (column “Data Access Regime (most relevant examples)”) below.

<sup>20</sup> See similarly the analysis of OECD (2019b, 24–48) Other inspiration for this taxonomy comes from the work of Graef, Husovec, and van den Boom (2019).

<sup>21</sup> For the difference between data and information, see note 1.



sector organisations seeking access form the private access domain, while the public sector bodies form the authority access domain.<sup>22</sup>

### 3.2 Degrees of Access

It should be clear by now that access is not a monolithic entity, but a diverse reality whose features closely depend on how access opportunities are distributed. In short, access takes place on different degrees. The degree of access refers to the number of actors entitled to seek and obtain access to data under a given legal regime. So, with a little imagination we can picture data access as a vertical line or a pyramid whose summit is open access. The latter means providing indiscriminate access to the public, and has proven to be the ideal status to ‘maximis[e] the benefits of data, in particular in environments characterised by high uncertainty, complexity and dynamic evolution such as climate change, urban development and health care research’ (OECD 2015, 191). Legislation on public-sector data<sup>23</sup> is a striking illustration in this respect. The other end of the spectrum is closed access, under which seekers cannot access data due to the particular meaning or the sensitiveness of information (Pagallo 2014, xxii). This configuration characterises material labelled as classified information or state secrets, which can be exceptionally obtained if a seeker is provided with a security clearance. Access to trade secrets is closed as well.<sup>24</sup> Between open and closed access stands a grey area, covered by legal regimes allowing access to a restricted number of seekers. These scenarios are informed by a differential degree of access. So, for instance, since principles of contract law, such as freedom of contract and autonomy, govern data exchanges amongst actors, contracting parties can freely restrict access of third parties as they wish.<sup>25</sup> Other examples can be found in some sector-specific regimes. In this sense, access to vehicle repair and maintenance information<sup>26</sup> and to

---

**22** “Authority domain” is employed instead of “public domain”, which is an IP notion. In personal data processing practices an additional access domain crystallises into some data subjects’ rights (i.e. the rights of access and data portability), which provide individuals with a legal position over information directly or indirectly identifying them. Personal access domain is not analysed in this contribution. See Table 1 below.

**23** 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 L172/56. Hereafter ‘Open Data Directive’.

**24** See Trade Secrets Directive, arts 3 and 4.

**25** See Section 2 above.

**26** Regulation (EU) 715/2007 of 20 June 2007 of the European Parliament and of the Council of 20 June 2007 on type approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6) and on access to vehicle repair and maintenance information [2007] OJ L171/1, art 6(1).

data about tests of chemical substances on animals<sup>27</sup> is granted to independent car repairers and to manufacturers or importers of a chemical substance respectively.

### 3.3 Access Approaches

Lastly and more importantly, any data access legal scheme follows an approach. An access approach is the legal and technological construct of the rules revealing access seekers' prerogatives. This criterion mingles a pure legal and coercive dimension, based on duties and rights, and a technological one, through which rules are embedded into the technological architecture of data processing systems (Lessig 2006). Four types of approach can be outlined: open data, joint access, mandatory access, and data portability.<sup>28</sup>

Open data denotes the most prominent legislative method for bolstering access. Non-discriminatory access and costs of access, free (re-)usage, and machine-readability are its key features. The open data approach notably inspires legislation on public-sector data, aiming to unleash the innovative potential of material such as maps, transport data, geospatial information, statistics data, and so forth. The Open Data Directive indeed rests on the principle that any access seeker can obtain and re-use (for commercial or non-commercial purposes) data held by public sector bodies.<sup>29</sup> Access is free of charge in principle, and in any event 'administrative charges should ... no longer exceed the marginal costs of making it available for re-use' (Commission 2015, 7), such as the expenses sustained to copy, provide and transfer documents, anonymise personal data and adopt measures to protect confidential information.<sup>30</sup> In addition, the Open Data Directive encourages public authorities to create and maintain "high-value datasets",<sup>31</sup> whose

---

**27** Regulation (EC) 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC [2006] OJ L396/1, arts 27 and 30.

**28** See similarly (OECD 2019b, 39). See Table 2 below.

**29** Open Data Directive, arts 3(1) and 4(1).

**30** Open Data Directive, art 6(1).

**31** High-value datasets concern these kinds of information: geospatial, Earth observation and environment, meteorological, statistics, companies and company ownership, and mobility (Open Data Directive, annex 1).

Table 1: A seeker-based taxonomy of data access.

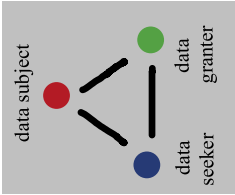
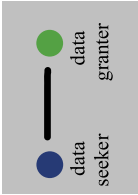
Nature	Access domain	Degree of access	Access approach	Data access regime (most relevant examples)
Personal data 	Personal access domain (seeker = data subject)	Differential	Data portability Mandatory access	<ul style="list-style-type: none"> <li>GDPR, art 20 (Right of Data Portability)</li> <li>GDPR, art 15 (Right of Access)</li> <li>Directive (EU) 2016/680 on data processing in criminal proceedings and investigations, art 14 (Right of Access)</li> </ul>
	Private access domain (seeker = private actor)	Differential	Joint access: Data markets Mandatory access	<ul style="list-style-type: none"> <li>Contract Law</li> <li>Essential Facilities Doctrine of competition law (“EFD”) (<i>controversial</i>)</li> <li>Directive (EU) 2015/2366 on payment services, arts 66 and 67</li> <li>Directive (EU) 2019/944 on metering and electrical consumption data, art 23</li> </ul>
	Authority access domain (seeker = public authority/institution)	Differential	Joint access: Data collaboration Mandatory access	<ul style="list-style-type: none"> <li>(Rules on) cooperation in (national) security, law enforcement, border control</li> <li>Criminal investigation law (law enforcement)</li> <li>Tax law</li> </ul>
Non-personal data 	Private access domain (seeker = private actor)	Differential	Data portability Joint access: Data markets Mandatory access	<ul style="list-style-type: none"> <li>FFDR, art 6 (Porting)</li> <li>Contract Law</li> <li>EFD (<i>controversial</i>)</li> <li>Regulation 1907/2006 on chemical information, arts 27 and 30</li> <li>Regulation (EU) 2018/858 on vehicle repair and maintenance information, art 61</li> </ul>

Table 1: (continued)

Nature	Access domain	Degree of access	Access approach	Data access regime (most relevant examples)
				– Directive (EU) 2019/944 on metering and electrical consumption data, art 23
		Open	Open data	– Open Data Directive
	Authority access domain (seeker = public authority/institution)	Differential	Joint Access: Data collaboration	– (Rules on) cooperation in (national) security, law enforcement, border control
			Mandatory access	– Criminal investigation law (law enforcement)
				– Tax law
		Open	Open data	– Open Data Directive

**Table 2:** Main features of the access approaches.

Access approach	Main features
Open data	<ul style="list-style-type: none"> <li>– = open and non-discriminatory access to data</li> <li>– Free reuse</li> <li>– Machine readability</li> </ul>
Joint access	<ul style="list-style-type: none"> <li>– = data exchange environment amongst 2+ actors</li> <li>– Two typologies: (i) data markets; (ii) data collaboration</li> </ul>
Mandatory access	<ul style="list-style-type: none"> <li>– = obligation to provide data access to an access seeker</li> <li>– Lack of technological qualification</li> </ul>
Data portability	<ul style="list-style-type: none"> <li>– = one-off (physical) copy of data into the facilities of an access seeker</li> <li>– Machine readability</li> </ul>

reuse benefits society, the environment and the economy<sup>32</sup> and supports the emergence of a high number of users.<sup>33</sup> As a default rule, public sector bodies must provide access to these datasets via Application Programming Interfaces (APIs), i.e. sets of computing protocols that facilitate software interaction and therefore promote a ‘smooth flow of data’ amongst the stakeholders (Borgogno and Colangelo 2019, 3). Despite its potential, the open data approach is currently not integrated into other pieces of legislation.

Joint access is another approach to data access. It rests on the creation of data exchange environments in which two or more entities participate. It can be divided into two subcases: data markets and data collaboration. The first is rooted in the general principles of contract law (i.e. freedom of contract) and inform most of data sharing activities amongst firms (Ottolia 2017, 221ff). Unlike other approaches, it implies consensual management and voluntary actions of both access granters and access seekers, and rest upon a differential degree of access. Instructive examples in this respect are data transfer agreements and data pools.<sup>34</sup> However, as shown in Section 2, access granters can constrain re-sharing activities by relying on this approach. On the other hand, data collaboration encompasses cases in which sharing of data (mostly amounting to semantic information)<sup>35</sup> takes place out of the markets. A poignant example is the informational mutual assistance and the

---

<sup>32</sup> Open Data Directive, art 2(11).

<sup>33</sup> Open Data Directive, art 14(2)(b).

<sup>34</sup> Data pools are environments where “firms agree to share their digitalised information regarding a given market, in reference to a given service or generally in an industry, or within an e-ecosystem” (Lundqvist 2018, 146).

<sup>35</sup> For the difference between data and information, see note 1.

creation of shared data management systems amongst public bodies at national and EU levels (Curtin and Brito Bastos 2020; Schneider 2014, 98–106).

Meanwhile, an access seeker can obtain data as a result of access granter's obligations to disclose it, upon the former's request. This legislative approach can be referred to as mandatory access. As with joint access, its degree of access is differential, since data is available only to several actors featuring specific characteristics, and not to the general public. However, it does not revolve around voluntary efforts of sharing data, but on non-consensual systems of data exchange. Under a mandatory access approach, access granters can generally 'regulate individual access by technical means since the data does not necessarily leave their platform' (Graef, Husovec, and van den Boom 2019, 18). It is the case of the obligations to grant access to automotive and chemical data under the relevant sector-specific legislation. Following the same logic, part of the literature has explored whether the essential facilities doctrine of competition law would require access granters to share data with any access seeker.<sup>36</sup> Another illustration of the mandatory approach concerns the authority access domain. Public bodies can generally request and obtain access to data (notably, information) for reasons of public interest or national security (for instance, in accordance with tax and criminal domestic laws), or in compliance with duties to inform other authorities (Schneider 2014, 101–2).

Lastly, data portability is an approach that is deemed a promising means to promote cross-sectoral re-use of data (OECD 2019b, 43). Its advantages and impact on society, however, are yet to be analysed.<sup>37</sup> Data portability follows a technology-oriented perspective, for it forces access granters to provide specific access seekers with access to data in structured, commonly used, and machine-readable formats. It consists in a one-off (physical) copy of data into the facilities of an access seeker having a special relation with a given access granter (Graef, Husovec, and van den Boom 2019, 18). Its most instructive application is Article 20 of the GDPR. Similarly, the FFDR,<sup>38</sup> which applies to non-personal data, provides for a self-regulatory mechanism based on codes of conduct at EU level listing best practices for data porting.<sup>39</sup> This latter regime, albeit innovative, may well nip in the bud its access-enabling potential since it is merely optional.

---

<sup>36</sup> See, Section 4 below.

<sup>37</sup> With respect to personal data, see (OECD 2019b, 44).

<sup>38</sup> 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 [2018] OJ L303/59. Hereafter 'FFDR'.

<sup>39</sup> FFDR, art 6(1)(a).

Against this diverse background, two observations can be made at this stage. Evidence suggests that data access regimes appear not to offset the trends of *de facto* data ownership and fail to provide a proper basis of access to non-personal data in large-scale environments. First, some shortcomings relate to the fragmentation of data access regimes, which target market failures affecting some industries or serve specific purposes (e.g. reasons of public interest national security, and so forth) (Drexel 2017, 287). In doing so, these laws generally provide sufficient access to particular kinds of (meaningful) information,<sup>40</sup> but do not back up data access beyond their limited scope. This impacts the public sector to a great extent. In default of a legal footing, public bodies are marginalised from exchanges of privately-held data, the consequence being that reuse of datasets does not meet overall societal needs (Commission 2020a, 6–8, 2020b). Similarly, the data markets approach, which elects freedom of contract as the default rule, leads to unequal bargaining power and exclusion of smaller players.<sup>41</sup> Second, mandatory access approaches mostly govern access to some kinds of information, and not to large sets made up of other types of data and information. Lastly, some restrictions concern open data and data portability approaches. The first, despite its potential, remains circumscribed to public-sector data, whereas portability of non-personal data is merely an optional solution which is envisaged in codes of conduct.

Within this context, the burgeoning legal literature on this topic has taken a closer look at an instrument that might provide a horizontal framework of access to large sets of (non-personal) data, i.e. the essential facilities doctrine of competition law.<sup>42</sup> Since the latter has been a matter of intense debate, it is worth examining its applicability to data-exploitation environments.

## 4 Data Access under Competition Law: Data as an Essential Facility?

Competition law might prove useful in tackling issues of data governance, since it stretches over industries which are not subject to targeted regulation. Competition analysis certainly helps to identify power-related problems surrounding data ownership.<sup>43</sup> Applying remedies of competition law to enhance data access is,

---

<sup>40</sup> On the notions of data and information, see note 1 above.

<sup>41</sup> See generally Section 2 and COM (2020) 66 final, 6–8.

<sup>42</sup> Analysis of other legislation on data access is out of scope of the present contribution. For an overview, see the thorough analysis of Graef, Husovec, and van den Boom (2019).

<sup>43</sup> See Section 2.

however, a different story. The essential facilities doctrine (“EFD”) could be viewed as a solution in this respect.

The EFD refers to cases where a monopolist refuses to grant access to goods or services. If these amount to an essential facility for other market players to create their own products, the dominant undertaking has an obligation to ‘share them with everyone asking for access, including competitors’ (Colangelo and Maggolino 2018, 2). The doctrine, which evolved in EU case law starting in the early 1990s, has been applied both to material infrastructures and incorporeal assets. Nonetheless, it has not been enforced in the past decade (Graef 2019, 1–6). The CJEU’s reluctance displays the difficult balancing of rights which the EFD requires: freedom of contract, right to select trading counterparties and freedom to property’s disposal on one hand, and right to access indispensable infrastructures on the other (Graef 2019, 6).

Fitness of the EFD for granting access to data needs to be assessed against a set of criteria. First, applying the EFD demands proof of the existence of both market dominance and an abuse, i.e. a conduct of the monopolistic entity substantively affecting the position of competitors. Second, according to the relevant EU case law, a refusal to grant access to a facility is deemed illegitimate if four conditions are met (the “exceptional circumstances test” or “ECT”).<sup>44</sup> The refusal is an abuse if it (i) refers to a good or service that is essential for conducting a business in a related (secondary) market; (ii) precludes effective competition in that market; (iii) thwarts the emergence of a new product for which there is consumer demand; and (iv) finds no objective justification. Yet, it is difficult to open up large data infrastructures by relying upon these yardsticks (Colangelo and Maggolino 2017, 19–26; Drexler et al. 2016, 9).

The various shortcomings of applying the EFD to large data infrastructures must be evaluated jointly. Under competition law, an undertaking proves to be a monopoly if it holds a resource which eliminates competition in the downstream markets. In short, competitors must not be able to do without it.<sup>45</sup> Moreover, as emerged in the *Microsoft* case,<sup>46</sup> market monopolisation does not necessarily stem from a (legally) exclusive position (e.g. where the dominant undertaking can rely upon IP protection), but may also derive from the availability to the monopoly of a resource that has become essential in practice (i.e. a standard). The relevant EU

---

<sup>44</sup> The CJEU has employed this test in the famous *Magill* case (Joined Cases C-241/91 P and C-242/91 P *Radio Telefis Eireann (RTE) and Independent Television Publications Ltd (ITP) v Commission* [1995] ECR I-743).

<sup>45</sup> As Advocate General Jacobs stated in the Oscar Bronner case, a duty to provide access is to be restricted to cases where refusals result in effective harm to competition (Case C-7/97 [1998] ECR I-07791, Opinion of AG Jacobs, para 61).

<sup>46</sup> Case T-201/04 *Microsoft Corp. v Commission* [2007] ECR II-3601.



case law has generally not found it hard to discern dominance in cases where an access seeker intends to access (semantic) information<sup>47</sup> that is the *sine qua non* of running a business in a given market (Drexel 2017, 281). In *Magill*<sup>48</sup>, *IMS Health*<sup>49</sup> and *Microsoft*<sup>50</sup> the CJEU held that the refusal to grant access to sole-source basic information (such as, respectively, TV show schedules, insights into regional pharmaceutical sales embedded into a database, and interoperability information) is an abuse of dominant position. It is more difficult, however, to draw similar conclusions when it comes to large data amounts. Showing their essentiality under the first condition of the ECT (i.e. the “essentiality criterion”) is a challenging task as datasets are substitutable assets. Data can be virtually found in myriad other datasets (Drexel 2017, 281) Open data<sup>51</sup> and data commercialised by information brokers are an instructive example in this respect. Since many actors collect this data and there exist numerous access points, it cannot be deemed indispensable (Borgogno and Colangelo 2019, 12; Drexel 2017, 281). Furthermore, even if access to other kinds of data is curtailed (e.g. user data of digital platforms, machine-generated data), these can hardly be viewed as indispensable assets. They would be so only if it was ‘technically, economically or legally impossible to find substitutes’ (Colangelo and Maggolino 2017, 25).

However, as Drexel points out, evaluating data substitutability is challenging in most cases, since ‘even the petitioner for access, such as a big data analyst, will often only have a vague understanding about the kind of data contained in the dataset and about which data will produce the most valuable new information based on observable correlations’ (Drexel 2017, 281).

The latter argument affects the third condition of the ECT (the “new-product rule”). Ricolfi maintains that competitors seeking access to datasets are usually unaware of the new product or service which they are going to develop with data unless they have accessed it (Borgogno and Colangelo 2019, 12; Colangelo and Maggolino 2017, 25–26; Ricolfi 2017, 222; n.d., 21). Moreover, public bodies seeking access to data (so, for instance, for public interest purposes) would be ruled out in cases where they do not conduct business activities ‘in the sense of the concept of an undertaking under EU competition law’ (Drexel 2017, 284).

Another cluster of challenges regards practicalities and administration of data access. Concerns revolve around the fact that applying the EFD to large datasets

---

47 On data and (semantic) information, see note 1 above.

48 *Magill* (n 83).

49 Case C-418/01 *IMS Health GmbH & Co. OHG v NDC Health GmbH & Co. KG* [2004] ECR I-5039.

50 *Microsoft* (n 86).

51 Here reference goes to data included in the scope of application of the Open Data Directive.

would overlook the technological dimension of the data flow.<sup>52</sup> Particularly, it is not clear to which part of data the dominant undertaking should grant access, nor how access is to be designed over time (for example, should access be provided in a one-off solution, continuously, in real time?) (Colangelo and Maggiolino 2017, 27; Gal 2017, 14). Other issues concern the terms and conditions of the (compulsory) licence on data: it is not certain whether monopolists should licence under fair, reasonable and non-discriminatory terms (“FRAND”), as in the case of standard essential patents (“SEPs”) (Colangelo and Maggiolino 2017, 26–29).

In sum, the EFD, as currently construed by the CJEU, is not a suitable instrument for opening up access to large-scale datasets. Such a conclusion has prompted some literature to propose ways of making the EFD fit for purpose (Crémer, De Montjoye, and Schweitzer 2019, 98–108; Graef 2019). Nonetheless, it appears rather difficult to re-jig instruments of competition law so that they can address data access requests. In fact, remedies of competition law are built as case-by-case and *ex post* interventions to a given set of market failures (Drexel 2017, 279; Drexel et al. 2016, 9), aimed to minimise negative effects of exclusionary and predatory behaviours (Ullrich 2019, 19). They rather “[react] *ex post* when data are (already) in the hands of only a few or dependencies are long established, and [will] never provide a sufficient basis for expanding access claims beyond the prevention of anti-competitive refusals to grant access” (Ullrich 2019, 30). This last argument suggests that competition law, albeit potentially heralding a shift towards a world where data is more accessible, is ill-suited for this job.

## 5 Raw Non-Personal Data and the Commons: A Closer Look

### 5.1 Another Limitation: From Non-Personal Data to Raw Non-Personal Data

As examined in the foregoing, recent technological progress has caused the rise of a *de facto* control over large-scale datasets. Existing data access regimes, however, are of limited relevance.<sup>53</sup> The findings of the preceding sections suggest that solutions for enhancing data access may lie elsewhere. Theories of the commons may be seen a way out of the predicament, for they map out institutional mechanisms of managing resources which centre on access and eschew exclusive control.

---

<sup>52</sup> See Section 3.

<sup>53</sup> See Section 2 and 3 respectively.

Before proceeding, it is necessary to set another limitation due to the restricted scope of this contribution. As is frequently the case of non-large scale processing environments, several types of non-personal data, being expressions of creative processes and amounting to (semantic) information, fall within the purview of intellectual property laws. It is the case, for instance, of the structured information of a marketing report, or data forming a music file. In such hypotheses, IP protection may be a reasonable access barrier, since it fulfils the goal of rewarding authors' ingenuie.<sup>54</sup> The same logic may apply to structured and processed data that goes through manipulations involving human intervention.<sup>55</sup> Bearing in mind these considerations, we need – at least for now – to relate the commons-based alternative to cases in which no (traditional) IP right backs up factual ownership of data.<sup>56</sup> In doing so, it is easier to balance the rights and interests over data and disregard any further analysis, albeit welcome, on the shortcomings of the traditional IP law system and the allocation of property rights. Accordingly, the scrutiny discussed below homes in on raw non-personal data (“RNPD”), meaning any non-personal that is defined by its representative characters (bits)<sup>57</sup> and has not undergone any automated or human activity of analysis, reuse or other manipulation aimed at extracting meaningful information from it or resulting from a creative effort.<sup>58</sup> This is particularly relevant in large-scale processing environments, where mass amounts of meaningless data elements contribute to form unstructured and non-relational (noSQL) databases.<sup>59</sup>

## 5.2 The Commons in Legal and Economic Scholarships

The commons as a conceptual category has caught the interest of numerous scholars for some time now. In particular, it has become a byword for anything that opposes predatory capitalism based on processes of extraction and power concentration (Marella 2017). Within the context of intangible goods, today's

---

54 Drahos points out that IPRs should be viewed as privileges created to fulfil predetermined goals and, as such, should be linked to holders' duties (Drahos 2016, 260ff).

55 In these hypotheses, data can eligible for database protection (either genuine copyright or *sui-generis* protection) under Directive 96/9/EC.

56 Factual ownership includes data protected as a trade secret, which is a quasi-IP tool. On this point, see Section 2.

57 See similarly the definition advanced by Zech (2016, 74).

58 On “raw data” and “processed data”, see the taxonomy in Commission (2020b, 93). Pre-processing and pre-selection which is purely aimed at making a dataset compatible with others do not amount to manipulation activities described here.

59 To use Gervais' wording, raw data form the basic material of ‘Big Data corpora’ (Gervais 2019).

capitalism has indeed fostered the emergence of a system of ubiquitous intellectual property rights. IPRs stretch to almost any immaterial resource stemming from the human mind (e.g. software, biotechnological inventions, and so forth) in the name of “economic coercion” (Braithwaite and Drahos 2000, 79) and the dominance of a few powerful actors (Mattei and Nader 2008, 83–88).

The theories of the commons stand in the way of the neo-liberal capitalistic mantra by suggesting a generative model of property (Mattei and Quarta 2018, 49–50) and foster models of resource management in which “freedom to operate outweighs power to appropriate” (Benkler 2014, 76). Literatures in economics and law, however, have struggled to attain consensus about an all-encompassing definition (Holder and Flessas 2008, 300; Marella 2017, 65; Mattei and Quarta 2018, 48). As maintained by the Ostrom school, the commons are social systems comprising three components: (i) common pool resources; (ii) a community that has access to and takes care of this resource; and (iii) a collective action of creating, restoring, maintaining, and governing in common (“commoning”) (Marella 2017, 66).<sup>60</sup> Several legal authors have then tried to broaden Ostrom’s perspective to regulate matters of information policy by moving forward the digital commons,<sup>61</sup> and, more recently, contractually-based computational commons (Ottolia 2017, 288ff). Other strands of legal literature, meanwhile, have sought to re-interpret property law in view of the commons as a way of putting use value first, rather than exchange value,<sup>62</sup> and incorporating humanitarian values. This view has breathed life into reforms of domestic legal systems, for example in Italy.<sup>63</sup>

The subsequent sub-sections investigate whether pigeonholing RNP as a commons can enhance access to it. Two main reasons substantiate why commons management of RNP may be desirable. First, RNP can be deemed a cooperative infrastructural resource that calls for being pulled out of its factual enclosure to open up the benefits of DDI to a greater number of actors.<sup>64</sup> This understanding is referred to as the “structuralist approach” in the remainder of the paper. Second, grasping RNP as a commons means valuing its functional nature, making data available to a wide number of actors for the fulfilment of fundamental rights and enhancing human flourishing. This latter perspective is therefore called “functionalist approach”.

---

**60** See in particular the thorough and classic analysis by Ostrom (1990).

**61** For a thorough analysis of the evolution of the commons theories in the US, see Benkler (2014, 74–77).

**62** This distinction have been drawn by Marx and put to use in the IP context in the mid 90s (Drahos 2016, 111ff). In the current debates on the commons, see the contributions of De Angelis (2017, 29) and Mattei and Quarta (2018, 31).

**63** See Section 5.4. below.

**64** Some initial remarks in this sense are brought forward by Sappa (2019, 416).

### 5.3 The Structuralist Approach: RNPDP as a Cooperative Infrastructure

Quite obviously, understanding RNPDP as a cooperative infrastructure means dividing the scrutiny into two steps. First, RNPDP is shown to be the result of cooperative processes. Its infrastructural dimension is examined thereafter.

The first component lies in empirical evidence – which is usually overlooked by legal scholarship – and has to do with the relational feature of data. In fact, RNPDP is produced as a result of the activities of different actors which are involved in a web of voluntary and (mostly) involuntary cooperative relations. This is the case of many aspects of human life (Benkler 2011). On closer inspection, RNPDP (and data in general) is a depiction of cooperative interactions without which its analysis and reuse would not be that valuable. There are many instructive examples in this respect. For instance, busses and trams equipped with automated sensors produce RNPDP on urban traffic which would not come into existence without passengers. Similarly, industrial RNPDP would not exist if the efforts of the factory workers activating or utilising smart machinery were disregarded. Data acquired by interconnected tractors and agricultural machinery likewise reflects farmers' activities in the countryside. In any application scenario, RNPDP is just a by-product or a side effect of activities in which many actors engage.

All these cases show that *de facto* data ownership obscures the inherent cooperative dimension of data production environments. Acknowledging factual dominium of data indeed depends on a short-sighted logic which means 'accepting private appropriation of what belongs to the public domain, i.e. the data' (Ullrich 2019, 27). On the contrary, commons management has proven to be desirable in respect to collaborative organisational models, such as peer production (Benkler 2016b, 110–14). Understanding RNPDP as a commons may therefore serve the purpose of enabling those who cooperate to access data.

Recognising the cooperative dimension of RNPDP is a promising starting point for applying commons-based mechanisms of resource management. Besides this, commons management fits RNPDP because of its infrastructural features. The examples above rather hint at it as well. According to one of the most thorough contributions on the topic, infrastructures are assets which meet the following criteria: '(1) The resource may be consumed nonrivalrously for some appreciable range of demand. (2) Social demand for the resource is driven primarily by downstream productive activities that require the resource as an input. (3) The resource may be used as an input into a wide range of goods and services, which may include private goods, public goods, and social goods' (Frischmann 2012, 61).

RNPD falls into such three requirements. First, nonrivalry means that consumption of a good by one individual does not affect or detract from simultaneous consumption opportunities of others. RNPD is a partially nonrival good, in the sense that it can be consumed and reused by an indiscriminate range of users for an unlimited number of times (OECD 2015, 181) but can be exhausted or depleted ‘at a rate that does not immediately transform the infrastructure but still may reduce its capacity and require maintenance or replenishment over time’ (Frischmann 2012, 63). Exhaustion and depletion mainly result from digital obsolescence, which may prevent older formats of datasets from being read by more advanced processing systems. In addition, RNPD is subject to depreciation, which occurs when it becomes irrelevant for the intended purposes of its processing (OECD 2015, 181).

The second factor connects infrastructures with productive activities. RNPD is a commodity that can be used as ‘an input into a wide range of goods and services’ (OECD 2015, 179) generating positive externalities from which society can benefit as a whole (Frischmann 2012, 112; Marella 2017, 72). In doing so, RNPD typifies a capital good, i.e. a kind of resource producing social benefits related to their downstream uses. RNPD, equal to other capital goods, is a means rather than an end (Frischmann 2012, 63–64). In fact, societal demand of RNPD is driven by demand for the outputs, i.e. the downstream activities enabled by RNPD analysis and reuse. RNPD collection by itself is pointless. RNPD is, on the contrary, the elementary material which, being processed through algorithms, generates information as an output in diverse scenarios.

Through the lens of the third criterion, RNPD can be viewed as a general-purpose input whose reuse engenders outputs in view of users’ ‘capabilities, options, opportunities, choices, freedoms’ (Frischmann 2012, 65). These outputs consist in private, public, and social goods. Social goods ‘generate value through their impact on social interdependencies and systems’, and include non-market goods (e.g. natural resources and ecosystems), merit goods (for instance, education, health-care), and social capital (for example, good will, fellowship, and sympathy) (Frischmann 2012, 43–48). This kind of goods has significant positive spill-overs which extend to society as a whole (Marella 2017, 72). General societal needs, however, are neglected if only private demand of RNPD is taken into account, merely reflecting users’ willingness to pay. In addition, the overall social value, albeit considerable, is particularly knotty to gauge. Problems in such measuring turn into a demand-manifestation problem affecting infrastructure allocation, design, investment, and management (Frischmann 2012, 66) and bring about ‘an optimization of the infrastructure design or prioritization of access and use of the infrastructure for a narrower range of uses than would be socially optimal’ (Frischmann 2012, 66). Regulators are aware of the general-purpose dimension of data. The Commission has illustrated how data plays a pivotal role in

serving social goods: large-scale data processing indeed enables to tackle today's central challenges (environmental degradation, climate change), improve health-care systems or confront emergencies (floods, wildfires) (Commission 2020a, 6–7). Nonetheless, as is the case for other infrastructures, data-driven production of social goods is stifled since data holders restrict availability of data they possess (Commission 2020a, 6–7).

General purposes of RNPd utilisation particularly emphasise the need for managing RNPd as a commons by valuing non-discriminatory and open-access regimes (OECD 2015, 182). If applied to infrastructures, commons management presents redistributive effects and stimulates ‘a virtuous circle between the spill-overs from certain uses and the social demand for access and social goods’ (Marella 2017, 73) (“cross-subsidisation”) (Frischmann 2012, 111–12). Yet, digital traits of RNPd call for particular forms and principles of governance differing from management systems of physical common-pool resources (Prainsack 2019, 7). Data-exploitation technologies, being prone to centralisation of control, do not spontaneously result in commons management,<sup>65</sup> but need a social construct which actively affects the relation between them and those controlling them (Brancaccio 2019, 863–64; Hardt and Negri 2017, 110). So, since they convert reality into a commodifiable quantity (Hess and Ostrom 2007, 10) which data holders store in their datacentres around the world, regulatory intervention may consist in imposing forms of commons management (Frischmann 2012, 110). Governance tools should therefore provide non-discriminatory access to RNPd to any access seeker, and exclude infrastructure owners (i.e. access granters) ‘from differentially allocating and prioritizing infrastructure access and use’ (Frischmann 2012, 110). To use the taxonomy of access regimes, the desirable default rule would then be the open data approach.<sup>66</sup>

In sum, the structuralist construct of the commons underscores the non-discriminatory and open component of access to RNPd. It sets aside ownership matters and lays solid foundations for data access. However, a purely legal positivist counterargument could be raised. The structuralist perspective by itself might be insufficient to implement and enforce the requests of access seekers. Data holders tend to not provide data access in absence of a well-shaped legal framework or incentives to do so (Commission 2020a, 6–8).<sup>67</sup> Embedding commons management into a right to data access may therefore help in this sense. To gain relevance, a legal position furnishing access seekers with access should be alert to

---

<sup>65</sup> Conversely, Rifkin argues that the Internet of Things easily leads to new forms of collaborative commons (Rifkin 2014).

<sup>66</sup> See Section 3.

<sup>67</sup> See also Sections 2 and 3.

specific rationales and justifications. This consideration paves the way towards a functionalist perspective, which merits further attention.

## 5.4 The Functionalist Approach: RNPd as a Values-Oriented Good

Allocation of a right to data access ought to be balanced against some overarching principles. The functionalist interpretation of the commons is germane to this case, as it values a humanitarian dimension.<sup>68</sup> In this guise, RNPd is a commons since it is a resource which, if accessed, allows a wide number of actors (individuals, communities, small and medium-sized companies, start-ups, public bodies and so forth) to fulfil fundamental rights, and enhances human flourishing.<sup>69</sup>

A right of access to RNPd would foster some fundamental rights at the core of the EU legal tradition. Notably, it would be consistent with freedom of expression and information as per Article 11 of the Charter of Fundamental Rights of the European Union (the “CFR”). As Hugenholtz puts it, these principles stand in the way of creating a new property(-like) right on data, for the notion of information encompasses syntactic data and comprehends commercial speech (Hugenholtz 2018, 66–67). Conversely, creating a right to data access may underpin fulfilment of freedom of information. Moreover, an access claim might enhance freedom of competition (Article 16 of the CFR) and freedom of services, which is one of the four freedoms of the EU Internal Market (Hugenholtz 2018, 67–69).<sup>70</sup> Viewed through this lens, access to RNPd would enable numerous actors to analyse and re-use it for their own purposes, regardless of any contractual limitation of reuse or re-sharing. They would be simply entitled to do so. Access granters would rather act as fiduciary entities demanded to provide data access in machine-readable formats

---

<sup>68</sup> Similarly, Drahos pioneered the idea of the need for an instrumentalist conception of IP which is oriented towards moral values (Drahos 2016, 231–65).

<sup>69</sup> This is the definition of the commons (or “common goods”, *beni comuni*) proposed in 2007 by the advisory panel in charge of reforming the Italian Civil Code in view of the commons (the *Commissione Rodotà*). Commons were understood as ‘those goods, publicly or privately owned, that are functional to the fulfilment of fundamental rights and individual flourishing and need to be protected by the law, also for the sake of future generations’. This category includes, *inter alia*, natural resources, cultural assets (either tangible or intangible), public services and infrastructures (Marella 2017, 67–68). Human flourishing is rather a complex concept which ‘is best understood with a reference to the teleological idea that everything that exists in life has an end’ (Sloot 2017, 266).

<sup>70</sup> Application of functionalist theories of the commons can thus be relevant *a contrario* in respect to Hugenholtz’s criticism of the building of a new data property right.



(Mezzanotte 2018, 527–28). Management costs of RNPDP would be shared (Mezzanotte 2018, 527–28).

The functionalist interpretation may appear to be too theoretical and detached from reality at first glance. Yet, this view has already proven crucial to take on board rights and interests of actors which are usually sidelined and not institutionalised. Urban environments are a vivid example in this regard. Cities have been the scene of capitalist accumulation under the aegis of the “smart city” label over recent years (Morozov and Bria 2018, 3). Commons management of RNPDP, however, have been successfully put into practice in some urban contexts through the creation of data access rules. Traditional top-down regulation has gradually left room for bottom-up tools promoting citizens’ participation to data production activities (Beckwith, Sherry and Par 2019, 205). So, for instance, Amsterdam and Barcelona are in the course of adopting decentralised data infrastructures to allot control over data harvested by contracting companies to the citizens (Morozov and Bria 2018, 31). Commons-friendly environments similarly make data gathered by public sensor networks available ‘for broader communal use’ (Morozov and Bria 2018, 32). As stated by Morozov and Bria, ‘a new cluster of start-ups, SMEs, NGOs, cooperatives, and local communities can take advantage of that data to build apps and services most relevant to them and the wider community’ (Morozov and Bria 2018, 32), since they have a right to do so.

## 6 Conclusion

This contribution has taken a first step towards depicting a management system of data based on access. Particularly, it has sought to describe the current trends of *de facto* ownership of data, capturing the state of the art in the literature which has evaluated it so far. In doing so, it has shown that factual control is based on access barriers that stifle data flow across society. The paper has then provided an overview of the legal regimes of data access, examining their differences against common criteria. As shown, the legislation on data access does not offset the shortcomings of factual data ownership, mostly because laws are targeted to address industry-specific market failures. The analysis has then progressed to scrutinising how the EFD can apply to data-exploitation scenarios as a way of fostering data access. However, as things stand the EFD shows manifold drawbacks in this respect.

Against this backdrop, theories of the commons can be viewed as a way out of the data-ownership puzzle. This analysis has not aimed to provide an exhaustive review of the literatures on the subject, but has explored the interactions between the commons and RNPDP. It is clear that the commons, taken both from a

structuralist and a functionalist perspective, open up golden opportunities to bolster access to RNPd and lay the foundations for a right to data access.

Numerous lines of research remain currently unexplored. First, implementing commons management in data environments may mean rethinking traditional IP protection in view of a more nuanced and flexible attitude. In this sense, some inspiration may come from the conception of ownership as a bundle of rights.<sup>71</sup> Trade secret laws will certainly require adjustments, which are worth scrutinising further. Further clarity should then be brought in respect to the design of commons management and the right to data access. As illustrated throughout the paper, empirical evidence suggests that the open data approach can work as a default rule governing RNPd. Future lines of work should elucidate, amongst other items, questions surrounding access remuneration (for example, should it be granted on FRAND terms?), the legal bases which underpin data access (e.g. public interest, legitimate interests of the access seeker?), and its technological configuration (i.e. implementation of interoperable formats?).

## References

- Beckwith, R., J. Sherry, and D. P. Part. 2019. "Data Flow in the Smart City: Open Data versus the Commons." In *The Hackable City: Digital Media and Collaborative City-Making in the Network Society*, edited by M. de Lange, and M. de Waal: Springer.
- Benkler, Y. 2011. *The Penguin and the Leviathan: The Triumph of Cooperation over Self-Interest*: Crown Business.
- Benkler, Y. 2014. "Between Spanish Huertas and the Open Road: A Tale of Two Commons?". In *Governing Knowledge Commons*, edited by B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 69–98: Oxford University Press, <https://doi.org/10.1093/acprof:oso/9780199972036.003.0003>.
- Benkler, Y. 2016a. "Degrees of Freedom, Dimensions of Power." *Daedalus* 145 (1): 18–32.
- Benkler, Y. 2016b. "Peer Production and Cooperation." In *Handbook on the Economics of the Internet*, edited by J. M. Bauer, and M. Latzer, 91–119: Edward Elgar.
- Borgogno, O., and G. Colangelo. 2019. "Data Sharing and Interoperability: Fostering Innovation and Competition through APIs." *Computer Law & Security Report* 35 (5): 1.
- Braithwaite, J., and P. Drahos. 2000. *Global Business Regulation*: Cambridge University Press, <https://doi.org/10.2307/3089030>.
- Brancaccio, F. 2019. "Appropriation, Common Property, the Inappropriable: Notes on the Law of the Common in Platform Capitalism." *South Atlantic Quarterly* 118 (4): 857–76.

---

<sup>71</sup> Reference goes to the famous theory pioneered by Honoré (1961). Ullrich underscores that the Anglo-American doctrine of the bundle of rights could help rejig IP in view of data environments (Ullrich 2019, 30).

- Colangelo, G., and M. Maggolino. 2017. "Big Data as Misleading Facilities." *European Competition Journal* 13 (2–3): 1–33.
- Colangelo, G., and M. Maggolino. 2018. *Data Access and AI: Antitrust vs. Regulation*. Paper for EU Commission's Project "Shaping Competition Policy in the Era of Digitisation.
- Commission. 2015. *Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources*.
- Commission. 2017. *Building a European Data Economy COM (2017) 9 final*.
- Commission, 2018. *Towards a Common European Data Space COM (2018) 232 final*.
- Commission, 2019. *Guidance on the Regulation on a Framework for the Free Flow of Non-personal Data in the European Union COM (2019) 250 final*.
- Commission, 2020a. *A European Strategy for Data COM (2020) 66 final*.
- Commission. 2020b. *Towards a European Strategy on Data Sharing for the Public Interest*, <https://doi.org/10.2759/406717>.
- Crémer, J., Y. D. Montjoye, and H. Schweitzer. 2019. *Competition Policy for the Digital Era*. European Commission Report.
- Crétois, P. 2015. "La propriété repensée par l'accès." *Revue Internationale de Droit Economique* 28 (3): 319–34.
- Curtin, D., and F. B. Bastos. 2020. "Interoperable Information Sharing and the Five Novel Frontiers of EU Governance: A Special Issue." *European Public Law* 26 (1): 59–70.
- De Angelis, M. 2017. *Omnia Sunt Communia: On the Commons and the Transformation to Postcapitalism*: Zed Books.
- De Mauro, A., M. Greco, and M. Grimaldi. 2016. "A Formal Definition of Big Data Based on Its Essential Features." *Library Review* 65 (3): 122–35.
- Drahos, P. 2016. *A Philosophy of Intellectual Property*, 2nd ed.: ANU eText, <https://doi.org/10.1108/intr.1998.17208daf.009>.
- Drexel, J. 2017. "Designing Competitive Markets for Industrial Data - between Propertisation and Access." *Journal of Intellectual Property, Information Technology and E-Commerce Law* 8: 257–92.
- Drexel, J., R. M. Hilty, L. Desautettes, F. Greiner, D. Kim, H. Richter, G. Surblytė, and K. Wiedemann. 2016. *Data Ownership and Access to Data: Position Statement of the Max Planck Institute for Innovation and Competition of 16 August 2016 on the Current European Debate*. Max Planck Institute for Innovation and Competition Research Paper No. 16-10.
- Drexel, J., R. M. Hilty, J. Globocnik, F. Greiner, D. Kim, P. R. Slowinski, G. Surblytė, A. Walz, and K. Wiedemann. 2017. *Position Statement of the Max Planck Institute for Innovation and Competition of 26 April 2017 on the European Commission's 'Public Consultation on Building the European Data Economy*. Max Planck Institute for Innovation and Competition Research Paper No. 17-08, <https://doi.org/10.2139/ssrn.2959924>.
- Farkas, T. J. 2017. "Data Created by the Internet of Things: The New Gold without Ownership?" *Revista La Propiedad Inmaterial* 23: 5–17.
- Floridi, L. 2010. *Information: A Very Short Introduction*: Oxford University Press.
- Floridi, L. 2014. *The Fourth Revolution. How the Infosphere is Reshaping Human Reality*: Oxford University Press.
- Frischmann, B. M. 2012. *Infrastructure: The Social Value of Shared Resources*: Oxford University Press.
- Gal, M. S. 2017. "Competition and Innovation in the Digital Environment." In *Concorrenza e comportamenti escludenti nei mercati dell'innovazione*, edited by G. Colangelo, and V. Falce: Il Mulino.

- Gervais, D. J. 2019. "Exploring the Interfaces between Big Data and Intellectual Property Law." *Journal of Intellectual Property, Information Technology and E-Commerce Law* 10: 3–19.
- Graef, I. 2019. *Rethinking the Essential Facilities Doctrine for the EU Digital Economy*. TILEC Discussion Paper No. 2019-028, <https://doi.org/10.2139/ssrn.3371457>.
- Graef, I., M. Husovec, and J. van den Boom. 2019. *Spill-Overs in Data Governance: The Relationship between the GDPR's Right to Data Portability and EU Sector-specific Data Access Regimes*. TILEC Discussion Paper No. 2019-005.
- Grossi, P. 1981. *An Alternative to Private Property: Collective Property in the Juridical Consciousness of the Nineteenth Century*. University of Chicago Press.
- Hardt, M., and A. Negri. 2017. *Assembly*. Oxford University Press.
- Hart, H. L. A. 1958. "Positivism and the Separation of Law and Morals." *Harvard Law Review* 71 (4): 593–629.
- Hess, C., and E. Ostrom. 2007. "Introduction: An Overview of the Knowledge Commons." In *Understanding Knowledge as a Commons*, edited by C. Hess, and E. Ostrom: MIT Press.
- Holder, J. B., and T. Flessas. 2008. "Emerging Commons." *Social & Legal Studies* 17 (3): 299–310.
- Honoré, A. M. 1961. "Ownership." In *Oxford Essays in Jurisprudence*, edited by A. G. Guest, 107–47: Oxford University Press.
- Hughenoltz, P. B. 2018. "Against 'Data Property'." In *Kritika: Essays on Intellectual Property*, edited by H. Ullrich, P. Drahos, and G. Ghidini, 48–71: Edward Elgar.
- Janal, R. 2017. "Fishing for an Agreement: Data Access and the Notion of Contract." In *Trading Data in the Digital Economy: Legal Concepts and Tools. Münster Colloquia on EU Law and the Digital Economy III*, edited by S. Lohsse, R. Schulze, and D. Staudenmayer: Nomos, <https://doi.org/10.5040/9781509921218.0020>.
- Kerber, W. 2016a. *A New (Intellectual) Property Right for Non-personal Data? An Economic Analysis*. Magks Paper No. 37-2016.
- Kerber, W. 2016b. "Governance of Data: Exclusive Property vs. Access." *IIC International Review of Intellectual Property and Competition Law* 47 (7): 759–62.
- Kerber, W. 2017. "Rights on Data: The EU Communication 'Building a European Data Economy' from an Economic Perspective." In *Trading Data in the Digital Economy: Legal Concepts and Tools. Münster Colloquia on EU Law and the Digital Economy III*, edited by S. Lohsse, R. Schulze, and D. Staudenmayer, 109–33: Nomos.
- Lessig, L. 2006. *Code: Version 2.0*: Basic Books.
- Lohsse, S., R. Schulze, and D. Staudenmayer. 2017. "Trading Data in the Digital Economy: Legal Concepts and Tools." In *Trading Data in the Digital Economy: Legal Concepts and Tools. Münster Colloquia on EU Law and the Digital Economy III*, edited by S. Lohsse, R. Schulze, and D. Staudenmayer, 13–24: Nomos.
- Lundqvist, B. 2018. "Competition and Data Pools." *Journal of European Consumer and Market Law* 7 (4): 146–54.
- Marella, M. R. 2017. "The Commons as a Legal Concept." *Law and Critique* 28 (1): 61–86.
- Mattei, U., and L. Nader. 2008. *Plunder: When the Rule of Law Is Illegal*: Blackwell Publishing.
- Mattei, U., and A. Quarta. 2018. *The Turning Point in Private Law: Ecology, Technology and the Commons*: Edward Elgar, <https://doi.org/10.4337/9781786435187>.
- Mayer-Schönberger, V., and K. Cukier. 2013. *Big Data: A Revolution that Will Transform How We Live, Work, and Think*: Houghton Mifflin Harcourt.
- Mezzanotte, F. 2017. "Access to Data: The Role of Consent and the Licensing Scheme." In *Trading Data in the Digital Economy: Legal Concepts and Tools. Münster Colloquia on EU Law and the Digital Economy III*, edited by S. Lohsse, R. Schulze, and D. Staudenmayer, 159–87: Nomos.

- Mezzanotte, F. 2018. “I poteri privati nell’odierno “diritto dello sviluppo economico.” *Politica del diritto* 3: 507–30.
- Morozov, E., and F. Bria. 2018. *Rethinking the Smart City: Democratizing Urban Technology*. Rosa Luxemburg Stiftung.
- OECD. 2015. *Data-Driven Innovation. Big Data for Growth and Well-Being*. Paris: OECD Publishing, <https://doi.org/10.1787/9789264229358-en>.
- OECD. 2019a. *Artificial Intelligence in Society. Artificial Intelligence in Society*. Paris: OECD Publishing, <https://doi.org/10.1787/eedfee77-en>.
- OECD. 2019b. *Enhancing Access to and Sharing of Data*: OECD Publishing, <https://doi.org/10.1787/276aaca8-en>.
- Ostrom, E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*: Cambridge University Press, <https://doi.org/10.4135/9781446200964.n32>.
- Ottolia, A. 2017. Big Data e innovazione computazionale: Giappichelli.
- Pagallo, U. 2014. *Il diritto nell’età dell’informazione. Il riposizionamento tecnologico degli ordinamenti giuridici tra complessità sociale, lotta per il potere e tutela dei diritti*. Giappichelli.
- Pagallo, U., M. Durante, and S. Monteleone, 2017. “What Is New with the Internet of Things in Privacy and Data Protection? Four Legal Challenges on Sharing and Control in IoT.” In *Data Protection and Privacy: (In) Visibilities and Infrastructures*, edited by R. Leenes, R. van Brakel, S. Gutwirth, and P. De Hert, 59–78: Springer.
- Prainsack, B. 2019. “Logged out: Ownership, Exclusion and Public Value in the Digital Data and Information Commons.” *Big Data and Society* 6 (1): 1–15.
- Purtova, N. 2015. “The Illusion of Personal Data as No One’s Property.” *Law, Innovation and Technology* 7 (1): 83–111.
- Ribot, J. C., and N. L. Peluso. 2003. “A Theory of Access.” *Rural Sociology* 68 (2): 153–81.
- Ricolfi, M. 2017. “IoT and the Ages of Antitrust.” *Concorrenza e Mercato* 1: 215–32.
- Ricolfi, M. n.d. *Il Futuro Della Proprietà Intellettuale Nella Società Algoritmica*: Giurisprudenza Italiana (Forthcoming).
- Rifkin, J. 2014. *The Zero Marginal Cost Society: The Internet of Things, the Collaborative Commons, and the Eclipse of Capitalism*: Palgrave Macmillan.
- Rubinfeld, D. L., and M. S. Gal. 2017. “Access Barriers to Big Data.” *Arizona Law Review* 59: 339.
- Sappa, C. 2019. “How Data Protection Fits with the Algorithmic Society via Two Intellectual Property Rights – a Comparative Analysis.” *Journal of Intellectual Property Law & Practice* 14 (5): 407–18.
- Sartor, G. 1996. *Intelligenza artificiale e diritto: Un’introduzione*. Giuffrè.
- Schneider, J. P. 2014. “Basic Structures of Information Management in the European Administrative Union.” *European Public Law* 20 (1): 89–106.
- Spencer, A. B. 2001. “Antitrust and the Information Age: Section 2 Monopolization Analyses in the New Economy.” *Harvard Law Review* 114: 1623–46.
- Surblytė, G. 2016. *Data as a Digital Resource*. Max Planck Institute for Innovation and Competition Research Paper No. 16-12.
- Surden, H. 2014. “Machine Learning and Law.” *Washington Law Review* 89: 87–115.
- Tusikov, N. 2019. “Precarious Ownership of the Internet of Things in the Age of Data.” In *Information, Technology and Control in a Changing World: Understanding Power Structures in the 21st Century*, edited by B. Haggart, K. Henne and N. Tusikov, 121–48: Palgrave Macmillan.

- Ullrich, H. 2019. *Technology Protection and Competition Policy for the Information Economy. From Property Rights for Competition to Competition without Proper Rights?*. Max Planck Institute for Innovation and Competition Research Paper No. 19-12.
- van der Sloot, B. 2017. "Privacy as Virtue: Searching for a New Privacy Paradigm in the Age of Big Data." In *Räume Und Kulturen Des Privaten*, edited by E. Beyvers, P. Helm, M. Hennig, C. Keckeis, I. Kreknin, and F. Püschel, 247–72: Springer, <https://doi.org/10.1007/978-3-658-14632-0>.
- Wiebe, A. 2017. "Protection of Industrial Data – a New Property Right for the Digital Economy?." *Journal of Intellectual Property Law & Practice* 12 (1): 62–71.
- Wiebe, A., and N. Schur. 2019. "Protection of Trade Secrets in a Data-Driven, Networked Environment - is the Update Already Out-Dated?." *Journal of Intellectual Property Law & Practice* 14 (10): 814–21.
- Zech, H. 2016. "Data as a Tradeable Commodity." In *European Contract Law and the Digital Single Market: The Implications of the Digital Revolution*, edited by A. De Franceschi, 51–80: Intersentia.