INTERNET POLICY REVIEW
Journal on internet regulation

**Volume** 10 **Issue** 3

RESEARCH ARTICLE

OPEN ACCESS

PEER REVIEWED

# Embedding European values in data governance: a case for public data commons

**Jan J. Zygmuntowski** *Kozminski University* jzygmuntowski@kozminski.edu.pl
**Laura Zoboli** *University of Warsaw***Paul F. Nemitz** *College of Europe*

**DOI:** https://doi.org/10.14763/2021.3.1572

**Published:** 30 September 2021
**Received:** 12 November 2020 **Accepted:** 4 February 2021

**Funding:** The article was partly supported by the National Science Centre, Poland (decision UMO-2018/31/B/HS5/01192. The authors did not receive any funding for this research.
**Competing Interests:** The author has declared that no competing interests exist that have influenced the text.


**Citation:** Zygmuntowski, J. J. & Zoboli, L. & Nemitz, P. F. (2021). Embedding European values in data governance: a case for public data commons. *Internet Policy Review*, *10*(3). https://doi.org/10.14763/2021.3.1572

**Keywords:** Data governance, Internet policy, Data commons, Public interest, Trust, European values

**Abstract:** The public sphere needs an 'ecosystem of trust' which could set out objectives of re-usage of data for the common good while protecting individual rights. This study analyses the emerging models of data governance through the lenses of science and technology studies (STS), critical data studies (CDS), and institutional economics, investigating which data governance model creates conditions for data stewardship guided by European values and rights. We critically examine two prominent, yet highly arguable, paradigms related to data, asserting that the systemic level of data assemblage must be re-conceptualised to reject the data-as-a-commodity view and take public interest into consideration. For data stewardship to achieve its goals, it is necessary to consider the inherent properties of data as commons, in the sense of a common-pool resources (CPR) framework. Therefore, we point towards public data commons as the model that is best suited to secure European rights and values while increasing data sharing at the same time. The design of such public data commons is the challenge of our time.

This paper is part of **Governing "European values" inside data flows**, a special issue of *Internet Policy Review* guest-edited by Kristina Irion, Mira Burri, Ans Kolk, Stefania Milan.

# 1. Introduction

Once only a grim reality envisioned by alarmists, the struggle over data has become a daily challenge and a leading topic in contemporary public debate. This struggle encompasses both personal data, which can be used to identify a person, as well as non-personal data [1]. The proposal of a digitally inclusive society [2] is under threat from unrestrained, highly disruptive developments in the fields of data extraction, algorithmic profiling, and the training of *artificial intelligence* (AI) on troves of *big data* (Isin & Ruppert, 2015). This new digital economy is described in a variety of ways: as a cognitive capitalism, where a systematic process of privatisation of information in different forms allows for maximisation of profits (Bauwens et al., 2019; Fumagalli et al., 2019), as a data colonialism, in which data is used to subjugate and transform social relations (Couldry & Mejias, 2019), or as a form of surveillance capitalism, in which various technological apparatuses monitor humans and try to predict and control human behaviour for the sake of profit (Zuboff, 2019). Regardless of the theoretical approach, the key assumption that current data mining practices are failing individuals and society remains at the heart of the problem.

A stark example of governance failures is the collaboration between Alphabet's DeepMind Health and Royal Free NHS Foundation Trust (Powles & Hodson, 2017; Winter & Davidson, 2019). Details about 1.6 million patients have been provided with inadequate efforts to secure data protection and usage clarity, thus resulting in a (legally ambiguous) violation of contextual integrity (Nissenbaum, 2010). Yet another consequence was significant value leakage to a private company, Google, which has become the largest collector of health and patient data worldwide. In this respect, one can speculate that Google bought DeepMind to get to the NHS data, and not just for its programming and development skills (Nemitz & Pfeffer, 2020).

On the other hand, an insufficient level of data sharing remains a key challenge for innovative organisations, especially when the data of private actors is examined in

1. In this paper, the term *data* includes both of these categories, with differentiation as necessary.
2. Understood as a society in which public interest guides usage of data and allows for the emergence of equal digital citizens. See, Isin & Rupper, 2015.

areas linked to provisioning of public services (Hofheinz & Osimo, 2017). According to Mitchell and Brynjolfsson (2017), new types of public-private partnerships should be installed, including tools to incentivise the collection of data for use in the public interest. The pooling of such data is still rare in public administrations or trustworthy intermediaries, which fails to capture value for public benefit. Similarly, the debate on how to "unlock private data for good" begins with creating conditions for access to private data, in order to achieve legitimate public policy goals (Alemanno, 2018).

Although efforts are made to disrupt the highly intrusive processes of data mining and datafication, these efforts have largely consisted of bottom-up activism, unlikely to scale and stabilise in an arrangement that leverages the digital economy for public purposes (Beraldo & Milan, 2019). Therefore, the public sphere needs an 'ecosystem of trust' (Mulgan & Straub, 2019), which could achieve objectives such as citizen empowerment, improvement of public welfare, or re-usage of data for the common good; all while safeguarding individual rights. Such governance is most notably the aim of the European Union (EU), seeking to align the development of technology with a set of core values constitutive for European society. Prior efforts to regulate transnational data flows exposed the insufficient scope of the European legal regime compared to Brussels' normative objectives. Hence the growing necessity for a comprehensive data governance framework, effectively enforcing European values and rights, where they apply, even beyond the EU territory.

In light of the context framed above, the key research question posed by this paper is: which data governance model creates conditions for data stewardship guided by European values and rights? The analysis is largely informed by science and technology studies (STS), critical data studies (CDS) and institutional economics, specifically common-pool resources (CPR) theory. A constitutionalist perspective has been adopted with regards to European rights and values.

The article begins with an overview of current challenges in the European legal system with regards to data. Subsequently, it presents four data governance models conceptualised through a data infrastructure lens and critically examines two prominent, yet highly arguable, paradigms related to data. The following section scrutinises the current understanding of data commons and their stewardship, arguing that public data commons are best positioned to secure European values and rights, while increasing data sharing. Given a strong legal mandate for such a model, whether it succeeds depends largely on mechanisms governing access/excludability and power asymmetries of participating actors. The conclusion summarises the results and opens space for further investigation.

The article aims to contribute not only to academic discussions, but also to contemporary policy considerations. It is especially relevant in the context of the European quest for technological sovereignty and considering the impact of data-driven technologies (such as AI) on human rights (Council of the European Union, 2020; European Commission, 2020a). The digital indeed has a constitutive role in the current phase of European integration, which the European Commission has reaffirmed with its Data Strategy (European Commission, 2020b) and the subsequent proposal of Data Governance Act (European Commission, 2020c). We invite scholars to study the conditions, tools and design principles which will aid in the establishment of common European data spaces and comprehensive data-sharing framework which serve the public interest.

# 2. European rights and values in the context of data

## 2.1. Constitutive values and rights of the European Union

Present primary and secondary law in the EU provide a rich texture of rights and values for data, as well as balances between them. Many of the past and present discussions on the ethics of AI or digital charters of rights seem to reinvent the wheel as to basic rights and values. It is sometimes conveniently ignored that in the constitutional and legal history of Europe the values and rights on which there is consensus have been identified and set down in a binding way, in processes which carry higher legitimacy than any of the present debates on ethics in AI. [3] Thus, the challenge today is not to re-invent catalogues of values and potential rights, but to actualise and make operational the rights and values which have already been laid down in the rich texture of European Law—from anti-discrimination acquis to data protection. [4]

However, European rights and values are not limited to individual rights. Indeed, the values of European constitutionalism, most notably the triad of human rights, rule of law, and democratic procedures, encompass collective rights and guarantees of institutions. Article 2 of the Treaty on EU also acknowledges values that refer to economic conditions and societal relations, such as equality, solidarity, and tolerance (European Union, 1992/2002). Both individual and collective rights, as

3. For example, compare initiatives such as the Charter of Digital Rights championed by the Portuguese Presidency in the Council of the EU to the Charter of Fundamental Rights of the European Union.

4. See the Presidency Conclusions on the Charter of Fundamental Rights in the context of Artificial Intelligence and Digital Change (Council of the European Union, 2020, p. 3): "We want to ensure that the design, development, deployment and use of new technologies uphold and promote our common values and the fundamental rights guaranteed by the EU Charter of Fundamental Rights".

well as institutions, erode in a regime where regulators fail to act in the public interest by enforcing the law as it stands (Nemitz, 2018). In this context, it is critical to understand how these rights and values are balanced within the data economy, with data protection as the benchmark.

The protection of personal data and, therefore, of data belonging to identified or identifiable individuals, has often been labelled *a priori* as overriding and this approach can be read as a consequence of the fact that an imbalance exists between the power of entities (whether business entities or not) that process and have access to large amounts of personal data, and the power of the individuals to whom that personal data belongs to control their information (European Union Agency for Fundamental Rights, 2018). However, it must be recognised that the fundamental rights to privacy and protection of personal data are not absolute rights and can be restricted by law under certain conditions. Both the European Court of Human Rights and the Court of Justice of the European Union (CJEU) have consistently held that a balance of the fundamental rights to privacy and protection of personal data with other rights is necessary in the application and interpretation of Article 8 ECHR and Articles 7 and 8 of the EU Charter of Fundamental Rights (Court of Justice of the European Union, 2008, 2011; European Court of Human Rights, 2012).

## 2.2. The challenges of balance and enforcement

A balance must be struck between these fundamental rights and other EU values, human rights, and public and private interests, to ensure basic rights such as social rights, freedom of expression, freedom of the press, or freedom of access to information. By way of example, the GDPR (General Data Protection Regulation) explicitly requires EU member states to reconcile by law the right to the protection of personal data with the right to freedom of expression and information, including for journalistic, academic, artistic, or literary purposes.

Specifically, it asks to "adopt legislative measures which lay down the exemptions and derogations necessary for the purpose of balancing those fundamental rights" (see Recital 153 GDPR). However, nothing stands in the way of the same legislator, in full respect of the Charter of Fundamental Rights and the Treaties, adjusting this balance in later legislation. In this context, the European Data Protection Supervisor (EDPS) has issued several opinions (Wiewiórowski, 2019), including on initiatives aimed at broadening the sharing of information for law enforcement purposes.

Furthermore, the GDPR states that the protection of personal data cannot be a justification for restricting the free movement of personal data within the Union (see Art. 1(3) GDPR). This also represents the result of a balancing act, although justified by the consistent level of data protection achieved in the EU internal market.

The CJEU, as a rule, does not consider the economic interest of a data controller to prevail over the interest of the individual data subject in data protection. It is worth mentioning for this purpose the *Google Spain* (Court of Justice of the European Union, 2014) case, where the interest of a private individual was considered to have a content about him de-indexed by Google, which prevailed over the economic interest that the operator of such an engine had in processing, and the interest of the public in general, given that neither the person nor the information in question were of public interest. In the *Manni* case (Court of Justice of the European Union, 2017), the CJEU had to balance the EU rules on data protection and Mr. Manni's commercial interest in removing information related to the bankruptcy of his former company from the commercial register. In this case, the CJEU established that the public interest in accessing the information prevailed over Manni's right to obtain the erasure of the data.

Although the terms for balancing rights and values in Europe predate the data economy, the shortcomings in the enforcement of existing individual and collective rights are today's biggest weakness in the data economy. The logic of data protection would be to enforce the rules with the strictest rigour against those who collect and process most personal data, given that the risk of non-compliance rises with the amount of data collected, and the complexity of processing involved. It is also plain that central tenants of today's data economy have been described as illegal in detail, without effective enforcement having been taken to reinstall the law. Today's reality is that the Data Protection Authorities (DPA) have not yet given themselves rational common parameters for prioritising their interventions, such as, for example, a principle to investigate first *ex officio* the collectors of the biggest amount of personal data, as with the amount of personal data collected the need for full compliance with GDPR becomes ever more important in the public interest.

In the efforts of shaping novel data ordering in the EU, more focus is therefore needed to create and sustain infrastructure for effective rights enforcement. It is evident now that even a small set of values may need many rights and institutions. Since there are already conflicting approaches to regulatory design in the European digital market (Ó Fathaigh & van Hoboken, 2019), it is necessary to critically examine the proposed data governance models in the light of properties of data

themselves.

# 3. Critical discussion of data governance models

## 3.1. Four models for data governance

The rising popularity of data governance is closely connected with the growing recognition of the value of data. Only a decade ago, data governance was understood as control over and the management of data (DAMA International, 2009), with an understanding that it is primarily an internal, company task, centred on coherence and clarity of information. More recent definitions of data governance incorporate numerous elements outside strict control, and tend to embrace a more holistic approach, defining data governance as a cross-functional framework, with specified rights, obligations, and formalised procedures for their application (Abraham et al., 2019). Thus, instead of guiding a single company or organisation, data governance encompasses the entirety of transnational and trans-organisational data flows, from the macro-level of nation-states to the micro-level of citizens.

The paradigm shift in governance from market to inter-organisational, networked organisations has been observed to be taking place across different types of economic resources [5]; with a dimension related to multi-stakeholder engagement (Jemielniak & Przegalińska, 2020). Although data silos formed by a few technology companies are still the prevalent mode of data governance, with the profit motive clearly being the dominant rationale in data flows, there are pioneering attempts at creating collaborative governance regimes with public interest in mind.

Following Micheli et al. (2020), we identify four distinct models of data governance encompassing the plurality of both proposed and piloted options [6]. These models are comprised of: 1) personal data sovereignty [7], where data owners control data as personal object and individually decide on selling their information, 2) data collaboratives, which encourage businesses to establish partnerships and exchange data, 3) data cooperatives, which empower individuals to voluntarily pool

5. Especially with the emergence of such trends as collaborative consumption, peer production, platform cooperativism and crowd sharing, as well as introduction of participatory governance schemes in municipalities across the world.

6. We adopt different terms, closer to existing literature than those found in Micheli et al. (2020). We use *data collaboratives* instead of *data sharing pools*, as found in e.g., Verhulst et al. (2020). Moreover, we opt for *public data commons* in place of *public data trusts* because we find that in terms of their properties data are commons; see further discussion.

7. The term *sovereignty* here denotes rather personal autonomy than sovereignty of a state actor. Self-sovereign identity is a notion traced back to libertarian and cypherpunk ideas of the early internet era, thus it refers to independence and self-determination outside the state (Barlow, 1996).

data for mutual benefit, 4) and public data commons, which rely on public actors to aggregate and steward citizen and commercial data (Micheli et al., 2020).

Despite their varying characteristics, the emerging models of data governance can be juxtaposed and assessed based on their main differences and similarities. The models outlined by Micheli et al. (2020), guided by analytical parameters (dimensions) inspired by Mulgan and Straub (2019), demonstrate that the dimensions of governance mechanisms and governance goals are actually derivatives of other two dimensions: specifically, stakeholder control and value allocation (see Table 1).

**TABLE 1:** Analytical matrix of data governance models

| | | **STAKEHOLDER CONTROL / GOVERNANCE MECHANISMS** | |
|---|---|---|---|
| | | **INDIVIDUAL / RIGHTS-BASED** | **INSTITUTIONAL / TRUST-BASED** |
| VALUE ALLOCATION / GOVERNANCE GOALS | PRIVATE / GROWTH-DRIVEN | Personal data sovereignty | Data collaboratives |
| | PUBLIC / WELFARE-DRIVEN | Data cooperatives | Public data commons |

Governance mechanisms are closely intertwined with the type of stakeholders engaged in decision-making (Borrás & Edler, 2014), as data subjects have different mechanisms for exercising control than data controllers. For example, legal instruments may secure individual rights and claims to power over data; GDPR rights being a prime example (Calzada, 2019). Data subjects therefore maintain individual control based on data-related rights, whereas data controllers establish trust-based mechanisms like contractual arrangements and processes (e.g., monitoring, verification of access), which rely on and support the credibility of the engaged institutions.

Since governance goals—the value-based objectives for creating governance schemes (Winter & Davidson, 2019)—find their realisation in the final distribution of gains from data sharing, the outcome is in fact an economic value allocation for the benefit of either a specific agent, or a broadly defined public interest. The public interest might, and often will, benefit the individual agent in the long-term, so the dialectical tension is rather again about balancing the order of values rather than choosing winners.

As data can be used multiple times and by varying actors, it may theoretically be

possible to combine models to some extent. One should also consider that the processing of data leads to the creation of new data (OECD, 2019) and to a better understanding of the original data holder or data producer of added value of new uses of data. This may in turn inspire a learning process allowing the original holder or producer of the data to make better use of the data, to his/her own benefit too.

## 3.2. Critique of two paradigms in data governance

Yet the four data governance models do not seem equally capable of fostering data sharing while securing important values and rights. To provide a more nuanced description, the complex socio-technical ecosystem in which data is produced, collected, and processed is examined. The analysis draws largely from STS and CDS, which investigate infrastructure not merely as a positivist and value-free, objective materialisation of scientific progress (Iliadis & Russo, 2016), but as heterogeneous, closely interlinked data assemblages (ecosystems), consisting of, 'technological, political, social, and economic apparatuses', governing production and circulation of data (Kitchin & Lauriault, 2014). The previous approach often tended to either be naive techno-determinism, characterised by Shoshana Zuboff (2015) as "the idea that technology must not be impeded if society is to prosper", or a particular agenda made opaque by pseudoscientific claims and declarations of ethical practices (Levy & Johns, 2016). However, the ubiquity of data in contemporary society requires going beyond some of the valid, yet narrow perspectives that dominated the landscape of the discussion until now. Two lines of critique seem especially promising and well-grounded in existing literature from various disciplines, encompassing STS, CDS and institutional economics.

### 3.2.1. Data-as-a-commodity

First, the idea that data are commodities is dubious, to say the very least. Ontologically speaking, data are digital reflections of the existing reality, not autonomous things-in-themselves (thinking in Kantian terms, existing independently of observation). Rather, data "are already embedded in the social", since increasingly "every aspect of the world is rendered in a new symbolic dimension" (Zuboff, 2015). Scholars are increasingly calling to attention that the commodification of data, "dematerializes it from the facts or processes or so-called natural person from which it is derived" (Käll, 2020, p. 3). The same principle holds for data collected from objects, spaces, or nature. Data, as mere technology-based recording of reality, may misrepresent, oversimplify, or distort it. Nevertheless, digital systems rely on input to function.

The rights-based approach currently fails to secure the integrity of both personal and non-personal data, and the data ecosystem that produces it. Certainly, a broader consideration cannot overlook the fact that every system requires reproducibility if it is to survive; especially human society and social reproduction (Federici, 2012). There needs to exist the industry or nature to produce industrial data and non-personal data, and often non-personal data is very reliant on social reproduction of workers operating machines, while nature data is reliant on being captured by economic or public actors.

By commodifying data, it is possible to distance it from individuals, processes, and material conditions required for its production, hence dis-embedding data from wider social relations and values (Jessop, 2007). That, in turn, allows control over fictitious property, and extraction of value for profit; critical to the continuous expansion of capital to new territories (Hardt & Negri, 2017). This concept was evident to Dan Schiller in 1988, when he assessed that in capitalist logic, the value of information "stems uniquely from its transformation into a commodity - a resource socially revalued and redefined through progressive historical application of wage labour and the market" (Schiller, 1988, p. 41).

The real subsumption of existing production and consumption in the market sphere, mediated by novel technologies, is closely followed by a colonising move to capture the data and value from spheres of life which remain non-commodified, starting with nature and progressing into the most inner lives of humans; their dreams, emotions, and experiences (Couldry & Mejias, 2019; Zuboff, 2019). It is Polanyi's "Great Transformation" occurring again, this time on digital platforms and led by cognitive capitalists to "rewire the flows of data and ultimately money and power" (Kenney et al., 2020). The rights-based approach aims to protect from the perils of data misuse, but it does not break from the fictitious commodity form of data.

Creating and enclosing virtual networks, digital spaces, and data producing technological apparatuses has serious consequences for social relations, as it fundamentally distorts and reshapes the needs of commercial success (Fuchs, 2016). Consequently, the governance goal of an empowered individual results in self-commodification, and the noble aim of sharing market conditions "effectively transforms the social context of what used to be a favour and turns it into something to be bought or sold" (Jemielniak & Przegalińska, 2020). Data governance founded on the premise of widespread, market-based treatment of data as a commodity could, "lead to another form of the tyranny of fine print" (Kerry & Morris Jr, 2019), where any online transaction would involve the sale of data (e.g., about the

transaction).

In the long term it is argued that in the presence of externalities which reveal relevant information about other data owners, there will typically be a depression of prices, as leakage or sales of data by some users allows for inferences about others (through statistical analysis/machine learning), thus diminishing the cost of further surveillance (Acemoglu et al., 2019). Data access and sharing for public interest are still marginal in the EU digital marketplace, and that should be incentivised (IDC & Open Evidence, 2017; Kerber, 2017). Open data represents a negligible segment in the data realm (Blackman & Forge, 2017). However, even an increase in competition and open data sharing would not redress the problem of the excessively low price and oversharing of data, and only the presence of a mediator (third-party authority) would improve efficiency.

This mediator (data intermediary) introduces mechanisms of trust in data governance. Instead of relying on personal discipline and consent, which in fact might be considered a "victimisation discourse" that individualises systemic problems (Janssen et al., 2020), institutional control bases on trust in the institution to represent data subjects' rights and aims. It is not our assertion that rights are unnecessary, but that for governance they are inadequate alone. Moving away from the practices of data-as-a-commodity requires complementary efforts to find systemic solutions not on an individual, but societal level.

### 3.2.2. Neglect of public interest

Second, public interest is a commonly missed part of the equation, with the struggle not between choices of data protection versus profit, but between different private allocations of value, and public, collective gains. Championing data protection has undoubtedly led to the establishment of individual rights to data in the EU, and increased social awareness of infringements committed by digital platforms (Rossi, 2018). However, data protection is insufficient as an overarching governance aim because of the negative consequences of different data governance models, as well as of positive network or spillover effects, which may be much wider than privacy alone.

There exist intersecting interests in personal data, many of which fall into the category of public interest, involving, for example, security, medical, financial, and environmental issues (Kerry & Morris Jr, 2019). When examining the relationship between personal data protection and public interest needs, some argue that the latter should always prevail. The GDPR provides for legal grounds and derogations from the protection of individual rights, under certain conditions, in the event of

the performance of a public interest task. It entrusts EU member states with the responsibility of determining several public interest strands, and laying down the legal basis for usage of personal data in the public interest. It is also very specific on occasion, for example, when it comes to health and the clear priority given to use of personal data, even without consent, to avoid pandemics like that of COVID-19 at present.

While interest limiting the right to data protection needs to be laid down in law for a balanced approach (Borgesius et al., 2015), the current non-uniform scenario, with different member states laying down public interest laws, hinders a unified approach at the EU level. The outcome is neither the uniformity of protection of fundamental rights in the EU, nor a good functioning of the internal market. Although not preventing the free flow of data within the EU internal market, such national variations surely do not help, rather they may sometimes push the data controller (being that entity determining the purposes and means of the processing) to confine the processing within the territory of a specific member state(s). However, one could deem a certain amount of pluralism in data governance as not negative *per se* or actually desirable if it reflects the inherent characteristics of the specific interests at stake. Consequently, in some cases, such as relating to data protection at the workplace and its relation to labour law, and on the relationship between the freedom of the press and data protection, the GDPR explicitly allows for a certain degree of variation, as an exception to the general rule of harmonisation being the core condition for both a functioning internal digital market and a level playing field for all as regards protection of fundamental rights.

Although the EU's position is to allow the processing of personal data in the public interest, considering the need to monitor the pandemic and its spread, the GDPR also points out that the member states have the autonomy to maintain or introduce national provisions to further specify the application of the rules governing the processing of personal data, when in the performance of a task carried out in the public interest.

In summary, it is clear in European law that important public interests can be placed above data protection. This is only possible provided that it is done in the form of parliamentary law, and in a way that is necessary in a democratic society as well as proportionate to the public interest pursued. There is always a need for a legal basis (EU or member state law) that precisely meets a public interest objective, that is proportionate to the legitimate aim being pursued, and that details specific provisions to adapt the application of the GDPR rules.

Moreover, it has become evident that even issues linked to data protection are not solely about data protection itself. As mentioned above, single players may have control over data sets with an enormous value, from a public policy goals point of view (Alemanno, 2018)—control without any duty to ensure access or sharing for public interest. Furthermore, in the case of dominant players, abusive conducts may be carried out through the exploitation of personal data collected. In this sense, one may recall the case involving Facebook in Germany—convicted by the *Bundeskartellamt* (federal antitrust office) of abuse of dominant position, as it made the use of its social networking service conditional on users granting broad permission to collect and process their personal data (*Bundeskartellamt*, 2019).

The protection of personal data may be seen as an obstacle to data sharing (Zoboli, 2020). First, the application of GDPR to any data sharing operation with mixed databases entails a direct obligation to secure GDPR rights in the process of data sharing, and substantially inhibits companies from starting data sharing initiatives. In practical terms, this means that, to share data, all the obligations of data controllers and processors, and the rights of data subjects, should be fulfilled. In practice, however, company data governance techniques are now so sophisticated and supported by tools of automation, which allow the precise separation of personal and non-personal data, and beyond that, the identification and follow up on rights and privileges granted or acquired for every piece of data.

Both the company that shares the data and the one that receives it must have a legal basis for the processing of personal data and must ensure an adequate level of security throughout the overall data set, in compliance with Article 5 (f) of the GDPR. In addition, data protection or data security might need to be invoked by controllers as means of restricting access for other stakeholders, with the aim of protecting an organisation's competitive advantage of asymmetrical access to data (Calo, 2017).

There are proposals to make access to data mandatory where it constitutes a barrier to market entry, while fully respecting the GDPR (BEUC, 2020). Acting in the public interest requires an intervention in the complex socio-technical ecosystem, shaping it to explicitly increase the scope and quality of public services acting in accordance with European values, and to reinforce welfare-driven data sharing (Zygmuntowski, 2018). The ambition to fuel AI development, and other data-based innovation, will remain unsatisfied if private interest of data controllers is prioritised over public interest; and this is true for both personal and non-personal data.

In this context, the concept of "reverse-PSI" plays a role, that is, granting public

sector bodies access to reuse privately held data (Poullet, 2020). This would thus be the mirror image of the Public Sector Information regulation—currently embodied in the Open Data Directive—which governs the access to and reuse of public data. A statement in this direction has been made by the Expert Group on Business-to-Government Data Sharing appointed by the Commission, whose report moves in the direction of defining a strategy to increase the sharing of data collected by the private sector for the benefit of public authorities (High-Level Expert Group on Business-to-Government Data Sharing, 2020).

Moreover, one of the key findings of the Furman Review on competition in digital markets is that *ex ante* rules bring advantages, such as clarifying procedures (Furman et al., 2019), which emphasise the need to expand the set of by-design principles. This is primarily because an *ex ante* regime specifies what can and cannot be done, what directly impacts the architecture of digital systems, leading to different design of data flows and governance mechanisms. Thus, the regulatory interest evolves from data access solutions (both horizontal and sectoral) to data governance systems with new institutions (Kerber, 2020). An example of such a governance tool is algorithmic impact assessment, proposed by various stakeholders to condition access to data and its algorithmic processing on creation of value aligned with public interest (Nemitz, 2018).

At the EU level, there is a shift towards an *ex ante* regulatory approach that seems to be underpinned by the EU Digital Package and, specifically, by the proposals for the Digital Services Act (DSA, European Commission, 2020d) and the Digital Markets Act (DMA, European Commission, 2020e) published last December 2020. More specifically, the DSA aims to introduce proportionate and balanced rules to better regulate the digital ecosystem, and its core is the provision of an innovative framework for transparency and liability for online platforms. Whereas the DMA addresses economic imbalances and unfair trade practices that may be implemented by online platforms acting as gatekeepers. This regulation is the consequence of the assumption that large online platforms can control important platform ecosystems in the digital economy, allowing them to leverage their advantages, including access to huge amounts of data (European Commission, 2020a).

In summary, the critique of both the individualistic, rights-based approach, and the critique of the growth-driven, strictly private value allocation, point to the necessity of institutional and trust-based mechanisms with a public interest focus, with the aim to enhance common welfare in practical concordance with the protection of the fundamental rights of individuals.

# 4. Towards public data commons

## 4.1. Characteristics of data commons

The systemic level of data assemblage must be re-conceptualised, putting more emphasis on institutional solutions with public interest in mind. For the purpose of governance, the most dire questions of value, goals, stakeholder control, and mechanisms of enforcement suggest that, "it is useful to think less about data as a commodity to be bought and sold, and more as a shared asset or common good" (Bass, 2020, n.p.). This claim can be interpreted two-fold, as Purtova (2021) points out: as a normative claim (commons vs siloes) or in the sense of a common-pool resources (CPR) framework. Discussing inherent properties of data and data assemblage, we opt for the latter.

There is renewed interest in studies on CPR, partly because the insights are applicable to novel challenges related to digital data (Bollier & Helfrich, 2012). The tensions between public value and appropriation have been deeply studied by Elinor Ostrom, who defined CPR as resources which have a high level of subtractability (rival consumption may threaten them), and a low level of excludability (difficulty in limiting consumption) (Ostrom, 1990). Although due to their replicative nature data would seem to be a pure public good (Mitchell & Brynjolfsson, 2017), it is increasingly viewed as commons that should be governed collectively, thus unlocking the different uses and values of data for different stakeholders while protecting their rights (Coyle et al., 2020).

When Hess and Ostrom discussed knowledge commons, they pointed to technological progress as the facilitator of the ability to exploit the commons. Their explanation is that the "ability to capture the previously uncapturable creates a fundamental change in the nature of the resource, with the resource being converted from a non-rivalrous, non-exclusionary public good into a common-pool resource that needs to be managed, monitored, and protected, to ensure sustainability and preservation" (Hess & Ostrom, 2007). What was once true for knowledge and the possibility to capture it (e.g., via intellectual property rights), became true as well for data with the expansion of surveillance apparatuses.

Subtractability is thus reinterpreted to mean not only depletion of physical resources but such conditions where overconsumption endangers sustainability of the commons. It emerges once the data ecosystem expands from the periphery of our socio-economic system, coming to its very fore. Capturing and processing data leads to far-reaching consequences for the very data subjects that produce data.

When rights are harmed by misuse, trust in digital systems declines. Similarly, overconsumption by dominant actors is no longer non-distortionary, since it affects the distribution of data-related gains (e.g., innovation surplus, business insights) and leads to market capture. Although data consumption is not a rival *per se*, the outcomes contribute to rivalry in a capitalist economy.

If the usage of data commons is uncontrolled, in particular if no stewardship values are provided within the governance model, it will ultimately result in problems such as reinstated enclosure of data, or disempowerment of individuals, *vis-à-vis* digital companies (Purtova, 2017). Those negative externalities diminish the general sustainability of the data ecosystem. Data commons does not physically deplete, but is "threatened by undersupply, inadequate legal frameworks, pollution, lack of quality or findability" (de Rosnay & Stalder, 2020). Purtova's (2021) claim that "one's "enjoyment" of data does not lead to its deterioration or depletion" does not consider far reaching consequences if the data processor enjoys the data in an unsustainable manner.

It is not collective management or sharing itself that constitutes data as CPR, but the nature of data. Contrary to the calls for "moving beyond one-size-fits-all solutions" (Carballa Smichowski, 2019), trial and error scrutiny of different data governance models allows for moving towards the solution that takes the characteristics of data into account.

## 4.2. Governance of data commons

Commons-based peer production mediated by digital technologies has been studied and championed for many years now as an alternative mode of production (Bauwens et al., 2019; Benkler, 2006; Kostakis & Bauwens, 2014), although the problems of long-term, sustainable governance, value allocation, and data flows, as well as risks associated with the power of corporate actors co-opting open access commons, have not been well considered (Bauwens & Niaros, 2017; Papadimitropoulos, 2018). Bodó (2020) claims that the successes of corporate, extractive practices are actually "the failure of the commons-based peer production movement, because it refused to think about value". Indeed, while Wikipedia is still being held out as a prime model of public interest related peer production, other ecosystems of peer production like Github, bought by Microsoft, Red Hat, bought by IBM, and SSRN, bought by Reed Elsevier (now RELX) have become part of the ever more private profit dominated digital ecosystem. Wikipedia, created with substantial funding from the big digital corporates and their staff, is spared, to maintain a fading historic dream, namely the claim that freedom from democracy made

legislative rules plus technology will lead to the best public interest solutions, embodied in its most poetic form in the Declaration of the Independence of Cyberspace by John Perry Barlow.

Digital commons scholars point out that for a commons to reproduce both its objects (data) and subjectivities (trust, mutual aid), it has to interact with the market system outside the commons value circuit itself (De Angelis, 2017). However, oversharing or leakage of data to for-profit companies increases the risk of value extraction and drives the negative feedback loop, where top talents are attracted only to those companies and innovation happens for the sake of a limited customer group (or shareholders only) instead of common good (Weber, 2017). The cost burden for accessing data commons must be pertinent to the possibilities of a given actor, and relevant to the true aims of data usage.

A fine line between stewarding CPR in public interest and governance models protecting the dominant power needs be recognised. Data protection, for instance, and rules of access to data must be addressed in a way that allows access for smaller and more agile entities as well, instead of creating a silo that again would conform to procedures of the biggest industry players. Similar to other CPR governance schemes, data commons are not synonymous with unrestricted, open access, but in order to remain sustainable they might be bound by specific rules in regard to who and for what aims to benefit from sharing (Hess & Ostrom, 2007). Designing rules for who, why and with what consequences gets excluded is thus central to considerations on data commons (Prainsack, 2019). Access and exclusion are intertwined here and designing both cannot be void of addressing power asymmetries which undermine sustainability.

For those reasons, the notion of data stewardship gained considerable attention lately, departing from a strictly technical understanding towards a definition that encompasses sustainable, responsible management oriented towards improving people's lives and securing public interest (Verhulst et al., 2020). Balancing embedded values with the ability to control data usage is difficult but carries the possibility of deployment of a repeatable framework utilising automated decision-making systems and protocols (O'Hara, 2019).

Successful cases of data sharing in research communities in the sciences were to a large extent governed according to CPR principles, in contrast with a few cases ridden with conflict (Fisher & Fortmann, 2010). Where data sharing caused considerable conflict between data producers and users, it was due to lack of mechanisms predicted by theory: unclear appropriation rules and boundaries, low impact on

collective-choice arrangements, and recognition of rights to organise. Web-based data commons repositories also employ a variety of control policies and tools, guided by commons theory (Eschenfelder & Johnson, 2014). Those cases show that stewarding data commons is both effective and sustainable; yet scaling such solutions means changing the scope from narrow, industry-limited to societal and thus public.

## 4.3. Designing public data commons

To solve the data governance conundrum, we argue for institutional, trust-based mechanisms and welfare-driven value allocation. Taking into account the characteristics of data commons, this requires an intervention to establish public data commons, defined as a trusted data sharing space established in the public interest. Primarily, this includes safeguarding (or even advancing) European values and rights by active participation of public actors in stewarding data.

Mariana Mazzucato rightly asks: "why (should) the public's data not be owned by a public repository that sells the data to the tech giants, rather than vice versa?" (Mazzucato, 2019, n.p.). Indeed, there is a growing body of literature calling for digital public infrastructure, such as public platforms or data spaces (Bass et al., 2018; Hall & Pesenti, 2017; Morozov & Bria, 2018), mindful of natural monopolies, and network effects on digital platforms (Belleflamme & Peitz, 2016). Public data commons unwind the corporate extraction of data by creating capacity in the public sector to shape socially beneficial AI, and regain technological sovereignty (Zygmuntowski, 2020). Enforcing interoperability might be one of the ways to encourage data sharing with the public data commons; several countries have already adopted obligatory data sharing mandates, departing from voluntary measures (Zoboli, 2020). In line with the reverse-PSI approach mentioned, for example, in France, private actors are obliged to open specific categories of data in their possession under certain conditions (See *Loi du 7 octobre 2016 pour une république numérique*, Art.17 and following). In particular, the common element of these categories of data is that they are of *public interest*and include, for instance, data generated in the context of procurement or commercial data for the development of official statistics.

It is however essential not to undermine the possibilities of data sharing, knowledge spillovers and re-utilisation of data by different societal actors. Instead of monopolising data for other means, the role of the public sector is to be the facilitator and custodian, setting out governance rules, objectives and enforcing them, while inviting or mandating different data ecosystem stakeholders and granting

them participatory mechanisms. Collaborative governance acts as a systemic feedback loop, preventing the dominant organisation—and this can be a public actor, albeit in the present will often be a private actor—from rigging the system so that it continues with "business as usual" at the expense of other stakeholders (Susha & Gil-Garcia, 2019).

Based on the considerations discussed in this article, such governance should be based on a set of principles such as European values and rights, embedding them to guide rules of access/excludability and discrete choices on architecture design. Data-related rights would find their way as functionalities of public data commons interfaces or procedures. Their execution would not be at the mercy of data controllers, but a public service initiated by the commons on behalf of the data subject. In this sense, public data commons resembles the proposals of data trusts (Hardinges et al., 2019; O'Hara, 2019), yet the scope and main governance goal is much more universal given the role of public administration. Those considerations are summarised in Table 2, where we also point to technological, legal, and institutional solutions to main challenges of public data commons. Figure 1 is the proposal for governance design of public data commons which leverages those findings.

**TABLE 2:** Overview of public data commons characteristics

| ISSUE | DESIGN CHOICE | KEY PROBLEM | OSTROM PRINCIPLES | POSSIBLE SOLUTIONS |
|---|---|---|---|---|
| PRIVACY | Data protection by design | Secure computing<br>Preventing data leakage | 1. Clearly defined boundaries | 'Move algorithm to data': Open Algorithms (OPAL) principles (Hardjono & Pentland, 2017)<br>Differential privacy<br>Federated storage & learning |
| VALUE | Public welfare-driven allocation | Benefits decoupled from public interest | 2. Congruence between benefits and costs | Funding mission-oriented innovation (Mazzucato, 2018)<br>Licences granting returns to the public |

"Ostrom principles" refer to Ostrom (1990).

| ISSUE | DESIGN CHOICE | KEY PROBLEM | OSTROM PRINCIPLES | POSSIBLE SOLUTIONS |
|---|---|---|---|---|
| CONTROL | Trust-based institutional mechanisms | Power asymmetry between stakeholders | 3. Collective-choice arrangements | Collaborative governance empowering data stakeholders<br>Consensus building for strategic decisions |
| ACCESS & EXCLUSION | Accountability to preserve data commons | Detecting malicious intent<br>Appropriate sanctions | 4. Monitoring<br>5. Graduated sanctions | Algorithmic impact assessment<br>Monitoring and auditing access<br>Graduated fining & denylisting violators |

"Ostrom principles" refer to Ostrom (1990).

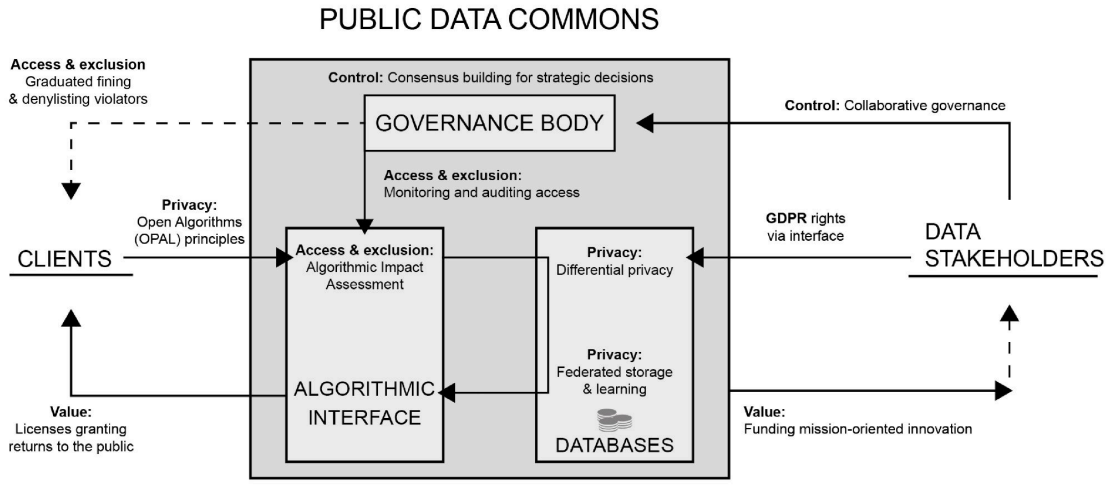


**FIGURE 1:** Governance design of public data commons

It is understandable that personal data sovereignty, data pods and self-determination leveraging data seems to champion digital emancipation. It is the promise of conscious, active citizens effectively bargaining on the markets while leveraging their rights to stay in control over data governance. However, data ownership means little if the owned property is sold on uneven terms, in a world of growing inequalities. Indeed, "decentralising processing does not necessarily imply decentralising power" (Janssen et al., 2020). To mitigate this potential problem, proposals of countermeasures have been developed, such as Jaron Lanier's data labour

unions (Arrieta-Ibarra et al., 2018), multi-client privacy administrators (Betkier, 2019), and consent champions (Ruhaak, 2020). This approach, combining personal data sovereignty with institutional mediation, still does not question the commodification of data and its social impact, but merely finds ways to secure more equitable distribution and control in a strictly market environment (Morozov, 2015).

Establishment of public data commons requires ambitious cooperation on the part of regulators, civil society, and businesses willing to benefit from the new data governance regime. It is unlikely that we can get only value allocation or stakeholder control right without drastic losses to the other. Inadequate effort in cooperating on value allocation to the public results in a deadlock between stakeholders lacking trust to set common standards, procedures, and exchanging valuable data, as it is observed currently (European Commission, 2018; Mitchell & Brynjolfsson, 2017). Conversely, too low cooperation to secure trust-based mechanisms for institutions prevents data cooperatives from scaling up to meaningfully boost universal welfare and offer a significant alternative to monopolistic data practices (Sandoval, 2020).

Public data commons are a model yet to be further explored, but one that might better inform how to structure common European data spaces (European Commission, 2018) or build a federated collaboration between national data repositories stewarded by EU member states. Similarly, the Data Governance Act aims to unlock the value of data while protecting individual rights. Unfortunately, these actions so far fall short of ambitions because they overemphasise private value creation and cling to the usual governance model of strictly market competition (between proposed data intermediaries). We believe that balancing value creation and data stewardship requires creating institutions, setting standards, and then inviting data stakeholders to collaboratively govern. This differs very much from the *ex post* logic of ordering existing actors.

Public data commons could act entrepreneurially, leveraging their role as digital utilities operators to ensure equal terms and protection of rights in data flows, but also serve as enablers of innovation and custodians of public interest performance monitoring. Such a scenario will not happen without sufficient public infrastructure, democratic oversight and expansion of governance using new legal tools and drawing from experiences of institution-setting of the past (Sadowski et al., 2021). This amounts to a systemic change that is more often called for; one that recognises the need to rejuvenate the European welfare state by thorough digitalisation, expansion of fundamental rights and increased collaboration. It remains to be hoped that the orientation for public interest outlined above will find its way into

a new set of European rules for good data governance, based on the rule of law, on fundamental rights and on the primacy of democracy.

## 5. Conclusion

In the search for a distinctively European understanding of technological sovereignty, encompassing both citizens' rights and technological artefacts that "tackle real (not only commercial) problems based on open codes" (Calzada, 2019, n.p.), there is a need for a more inclusive, broad governance design, addressing data across sectors and regulatory instruments. Given the fluidity of data, its replicability, and its own multipurpose nature, the thinking of the past will not serve public policy purposes anymore. If the same data can be used at the same time for a wide variety of purposes, both private and public, only a systematic view of all elements of governance will make optimised policy-making possible. Treating data as mere commodities or neglecting public interest has severe shortcomings, so the data governance of the future must consider that data are digital commons produced in an ecosystem that requires sustainability to thrive.

This highly complex system will only work with an elaborate set of rules, without which an increased concentration of power and a rise of societal disasters, brought about by current data extraction, will be seen. We argue that public data commons which leverage multi-stakeholder, collaborative governance policies, will increase data sharing, while safeguarding European rights and values, in the triangle between individual rights, economic growth and innovation, and public interest, as enshrined in European Primary and Secondary law.

We hope that more scholars will join in the effort to study the design principles, tools, and conditions for data commons to be stewarded in the public interest, since it is the challenge of our time. It is a natural and necessary extension of existing legal protections in the digital age, establishing institutions for European society venturing into the XXIst century.

## References

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, *49*, 424–438. https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Acemoglu, D., Makhdoumi, A., Malekian, A., & Ozdaglar, A. (2019). *Too much data: Prices and inefficiencies in data markets* (Working Paper No. 26296; NBER Working Paper Series). National

Bureau of Economic Research. https://doi.org/10.3386/w26296

Alemanno, A. (2018). Big data for good: Unlocking privately-held data to the benefit of the many. *European Journal of Risk Regulation*, *9*(2), 183–191. https://doi.org/10.1017/err.2018.34

Arrieta-Ibarra, I., Goff, L., Jiménez-Hernández, D., Lanier, J., & Weyl, E. G. (2018). Should we treat data as labor? Moving beyond 'Free'. *AEA Papers and Proceedings*, *108*, 38–42. https://doi.org/10.1257/pandp.20181003

Asociación nacional de establecimientos financieros de crédito (ASNEF) and federación de comercio electrónico y marketing directo (FECEMD) v. Administración del estado (joined case), (Court of Justice of the European Union 2011).

Barlow, J. P. (1996). *A declaration of the independence of cyberspace*. https://www.eff.org/cyberspace-independence

Bass, T. (2020). *It's time to think about our data as a common good*. British Council. https://www.britishcouncil.org/anyone-anywhere/explore/communities-connections/rethinking-data

Bass, T., Sutherland, E., & Symons, T. (2018). *Reclaiming the smart city: Personal data, trust and the new commons* [Report]. Nesta. https://www.nesta.org.uk/report/reclaiming-smart-city-personal-data-trust-and-new-commons/

Bauwens, M., Kostakis, V., & Pazaitis, A. (2019). *Peer to peer: The commons manifesto*. University of Westminster Press.

Bauwens, M., & Niaros, V. (2017). *Value in the commons economy: Developments in open and contributory value accounting*. P2P Foundation.

Becker, R., Thorogood, A., Ordish, J., & Beauvais, M. J. S. (2020). COVID-19 research: Navigating the european general data protection regulation. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3593579

Belleflamme, P., & Peitz, M. (2016). *Platforms and network effects Platforms and network e¤ects* (Working Paper No. 16–14). University of Mannheim, Department of Economics; RePEc IDEAS. https://ideas.repec.org/p/mnh/wpaper/41306.html

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom* (p. 515). Yale University Press.

Beraldo, D., & Milan, S. (2019). From data politics to the contentious politics of data. *Big Data & Society*, *6*(2). https://doi.org/10.1177/2053951719885967

Betkier, M. (2019). *Privacy online, law and the effective regulation of online services* (1st ed.). Intersentia. https://intersentia.com/en/effective-privacy-management-for-internet-services.html

BEUC. (2020). *Digital services act (ex ante rules) and new competition tool: Response to public consultations* [Position paper]. BEUC. The European Consumer Organisation. https://www.beuc.eu/publications/digital-services-act-ex-ante-rules-and-new-competition-tool-response-consultations/html

Blackman, C., & Forge, S. (2017). *Data Flows – Future Scenarios* [In-depth analysis / research report]. Centre for European Policy Studies. https://www.ceps.eu/ceps-publications/data-flows-future-scenarios/

Bodó, B. (2019). Was the Open Knowledge Commons Idea a Curse in Disguise? – Towards Sovereign

Institutions of Knowledge. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3502119

Bollier, D., & Helfrich, S. (Eds.). (2012). *The wealth of the commons: A world beyond market and state* (p. 442). Levellers Press. http://wealthofthecommons.org/

Borgesius, F. Z., Gray, J., & van Eechoud, M. (2015). Open data, privacy, and fair information principles. *Berkeley Technology Law Journal*, *30*(3), 2073–2131. https://doi.org/10.15779/Z389S18

Borrás, S., & Edler, J. (2014). The Governance of change in socio-technical and innovation systems: Three pillars for a conceptual framework. In S. Borrás & J. Edler (Eds.), *Borrás, s, edler, j (eds), the governance of socio-technical systems: Explaining change* (pp. 23–48). Edward Elgar Publishing.

Bundeskartellamt. (2019). *Decision B6-22/16*. https://www.bundeskartellamt.de/SharedDocs/Entscheidung/EN/Entscheidungen/Missbrauchsaufsicht/2019/B6-22-16.pdf?__blob=publicationFile&v=5

Calo, R. (2017). Artificial intelligence policy: A roadmap. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3015350

Calzada, I. (2019). Technological sovereignty: Protecting citizens' digital rights in the AI-driven and post-GDPR algorithmic and city-regional european realm. *Regions*. https://doi.org/10.1080/13673882.2018.00001038

Camera di commercio, industria, artigianato e agricoltura di lecce v. Salvatore manni, C-398/15 (Court of Justice of the European Union 2017).

Carballa Smichowski, B. (2019). Alternative data governance models: Moving beyond one-size-fits-all solutions. *Intereconomics*, *54*(4), 222–227. https://doi.org/10.1007/s10272-019-0828-x

Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism* (p. 323). Stanford University Press.

Council of the European Union. (2020). *Presidency conclusions—The charter of fundamental rights in the context of artificial intelligence and digital change* (Note 11481/20 FREMP 87 JAI 776). Council of the European Union. https://www.consilium.europa.eu/media/46496/st11481-en20.pdf

Coyle, D., Diepeveen, S., Wdowin, J., Kay, L., & Tennison, J. (2020). *The value of data: Policy implications* [Report]. Benett Institute for Public Policy, Cambridge; Open Data Institute. https://www.bennettinstitute.cam.ac.uk/media/uploads/files/Value_of_data_Policy_Implications_Report_26_Feb_ok4noWn.pdf

DAMA International. (2009). *The DAMA guide to the data management body of knowledge*. Technics Publications.

De Angelis, M. (2017). *Omnia Sunt Communia: On the commons and the transformation to postcapitalism* (p. 436). Zed Books.

Diekert, F. K. (2012). The tragedy of the commons from a game-theoretic perspective. *Sustainability*, *4*(8), 1776–1786. https://doi.org/10.3390/su4081776

Dulong de Rosnay, M., & Stalder, F. (2020). Digital commons. *Internet Policy Review*, *9*(4). https://doi.org/10.14763/2020.4.1530

Eschenfelder, K. R., & Johnson, A. (2014). Managing the data commons: Controlled sharing of scholarly data. *Journal of the Association for Information Science and Technology*, *65*(9), 1757–1774. https://doi.org/10.1002/asi.23086

European Commission. (2018). *Towards a common european data space* (COM(2018) 232 final).

European Commission. https://ec.europa.eu/digital-single-market/en/news/public-consultation-building-european-data-economy

European Commission. (2020a). *Digital Services Act package: Ex ante regulatory instrument for large online platforms with significant network effects acting as gate-keepers in the European Union's internal market* (Inception Impact Assessment Ares(2020) 287 7647).

*Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act)*, (2020) (testimony of European Commission).

Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on contestable and fair markets in the digital sector (Digital Markets Act), COM(2020) 842 final (2020). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0842&from=en

European Commission. (2020b). *A European strategy for Data*. European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1593073685620&uri=CELEX%3A52020DC0066

European Commission. (2020c). *Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act) (COM/2020/767 final)*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0767

Treaty on European Union (Consolidated Version), Treaty of Maastricht, C 325/5 Official Journal of the European Communities (2002).

European Union Agency for Fundamental Rights. (2018). *Handbook on European data protection law*. Publications Office of the European Union.

Federici, S. (2012). *Revolution at point zero: Housework, reproduction, and feminist struggle*. PM Press.

Fisher, J. B., & Fortmann, L. (2010). Governing the data commons: Policy, practice, and the advancement of science. *Information and Management*, *47*(4), 237–245. https://doi.org/10.1016/j.im.2010.04.001

Frischmann, B. (2012). *Infrastructure: The social value of shared resources*. Oxford Univesity Press. https://doi.org/ 10.1093/acprof:oso/9780199895656.001.0001

Fuchs, C. (2016). *Critical theory of communication*. University of Westminster Press.

Fumagalli, A., Giuliani, A., Lucarelli, S., Vercellone, C., Dughera, S., & Negri, A. (2019). *Cognitive capitalism, welfare and labour: The commonfare hypothesis*. Routledge. https://doi.org/10.4324/9781315623320

Furman, J., Coyle, D., Fletcher, A., Marsden, P., & McAuley, D. (2019). *Unlocking digital competition. Report of the digital competition expert panel* [Report]. HM Treasury.

Google spain SL, google inc. V. Agencia española de protección de datos (AEPD), C-131/12 (Court of Justice of the European Union 2014).

Hall, W., & Pesenti, J. (2017). *Growing the artificial intelligence industry in the UK*. Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy.

Hardinges, J., Wells, P., Blandford, A., Tennison, J., & Scott, A. (2019). *Data Trusts: Lessons from Three Pilots* [Report]. Open Data Institute. https://theodi.org/article/odi-data-trusts-report/

Hardjono, T., & Pentland, S. (2017, October 24). *Open Algorithms for Identity Federation* [Preprint]. ArXiv:1705.10880 [Cs]. http://arxiv.org/abs/1705.10880

Hardt, M., & Negri, A. (2017). *Assembly*. Oxford University Press.

Hess, C., & Ostrom, E. (2007). *A framework for analyzing the knowledge commons* (pp. 41–81). https://ieeexplore.ieee.org/servlet/opac?bknumber=6267279

High-Level Expert Group on Business-to-Government Data Sharing. (2020). *Towards a European strategy on business-to-government data sharing for the public interest* [Final report]. European Union. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=64954

Hofheinz, P., & Osimo, D. (2017). *Making Europe a Data Economy: A new framework for free movement of data in the digital age* [Policy Brief]. The Lisbon Council. https://lisboncouncil.net/wp-content/uploads/2020/08/LISBON-COUNCIL-Making-Europe-A-Data-Economy.pdf

IDC & Open Evidence. (2017). *Final report: European data market* (Study SMART 2013/0063). European Commission (Directorate-General for Communications Networks, Content and Technology). https://digital-strategy.ec.europa.eu/en/library/final-results-european-data-market-study-measuring-size-and-trends-eu-data-economy

Iliadis, A., & Russo, F. (2016). Critical data studies: An introduction. *Big Data & Society*, *3*(2). https://doi.org/10.1177/2053951716674238

Isin, E. F., & Ruppert, E. (2015). *Being Digital Citizens*. Rowman & Littlefield.

Janssen, H., Cobbe, J., & Singh, J. (2020). Personal information management systems: A user-centric privacy utopia? *Internet Policy Review*, *9*(4). https://doi.org/10.14763/2020.4.1536

Jemielniak, D., & Przegalinska, A. (2020). *Collaborative society* (p. 256). MIT Press.

Jessop, B. (2007). Knowledge as a fictitious commodity: Insights and limits of a polanyian perspective. In A. Buğra & K. Ağartan (Eds.), *Reading karl polanyi for the twenty-first century* (pp. 115–133). Palgrave Macmillan US. https://doi.org/10.1057/9780230607187_7

Käll, J. (2020). The materiality of data as property. *Harvard International Law Journal*, *61*. https://harvardilj.org/2020/04/the-materiality-of-data-as-property/

Kenney, M., Zysman, J., & Bearson, D. (2020). What polanyi teaches us about the platform economy and structural change. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3678967

Kerber, W. (2017). Rights on data: The EU communication "Building a european data economy" from an economic perspective. In S. Lohsse (Ed.), *Trading data in the digital economy: Legal concepts and tools*. Hart Publishing.

Kerber, W. (2020). From (horizontal and sectoral) data access solutions towards data governance systems. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3681263

Kerry, C. F., & Morris Jr, J. B. (2019, June 26). Why data ownership is the wrong approach to protecting privacy [Blog post]. *Brookings Techtank*. https://www.brookings.edu/blog/techtank/2019/06/26/why-data-ownership-is-the-wrong-approach-to-protecting-privacy/

Kitchin, R., & Lauriault, T. P. (2014). *Towards critical data studies: Charting and unpacking data assemblages and their work* (Working Paper No. 2; The Programmable City). Maynooth University.

Kostakis, V., & Bauwens, M. (2014). *Network society and future scenarios for a collaborative economy* (p. 87). Springer. https://doi.org/10.1057/9781137406897

Kostyuk, N. (2015, February). The digital prisoner's dilemma: Challenges and opportunities for cooperation. *2013 World Cyberspace Cooperation Summit IV, WCC4 2013*. https://doi.org/10.1109/WC

S.2013.7050508

Levy, K. E., & Johns, D. M. (2016). When open data is a Trojan Horse: The weaponization of transparency in science and governance. *Big Data & Society*, *3*(1). https://doi.org/10.1177/2053951715621568

Ma, Y., Lan, J., Thornton, T., Mangalagiu, D., & Zhu, D. (2018). Challenges of collaborative governance in the sharing economy: The case of free-floating bike sharing in Shanghai. *Journal of Cleaner Production*, *197*, 356–365. https://doi.org/10.1016/j.jclepro.2018.06.213

Mazzucato, M. (2018). Mission-oriented innovation policies: Challenges and opportunities. *Industrial and Corporate Change*, *27*(5), 803–815. https://doi.org/10.1093/icc/dty034

Mazzucato, M. (2019). *The value of everything: Making and taking in the global economy*. Penguin.

Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. *Big Data & Society*, *7*(2). https://doi.org/10.1177/2053951720948087

Mitchell, T., & Brynjolfsson, E. (2017). Track how technology is transforming work. *Nature*, *544*(7650), 290–292. https://doi.org/10.1038/544290a

Morozov, E. (2015, January). Socialize the Data Centres! *New Left Review*, *91*. https://newleftreview.org/issues/II91/articles/evgeny-morozov-socialize-the-data-centres

Morozov, E., & Bria, F. (2018). *Rethinking the smart city: Democratizing urban technology* (No. 5; City Series). Rosa Luxemburg Stifung, New York Office. https://www.rosalux.de/fileadmin/rls_uploads/pdfs/sonst_publikationen/rethinking_the_smart_city.pdf

Mulgan, G., & Straub, V. (2019). *The new ecosystem of trust: How data trusts, collaboratives and coops can help govern data for the maximum public benefit* [Paper]. Nesta. https://www.nesta.org.uk/blog/new-ecosystem-trust/

Nemitz, P. (2018). Constitutional democracy and technology in the age of artificial intelligence. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2133). https://doi.org/10.1098/rsta.2018.0089

Nemitz, P., & Pfeffer, M. (2020). *Prinzip mensch: Macht, freiheit und demokratie im zeitalter der künstlichen intelligenz*. Dietz Verlag.

Nissenbaum, H. (2010). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.

Ó Fathaigh, R., & van Hoboken, J. (2019). European Regulation of Smartphone Ecosystems. *European Data Protection Law Review*, *5*(4), 476–491. https://doi.org/10.21552/edpl/2019/4/6

OECD. (2019). *Going digital: Shaping policies, improving lives*. OECD Publishing. https://doi.org/10.1787/9789264312012-en

O'Hara, K. (2019). *Data trusts. Ethics, architecture and governance for trustworthy data stewardship* (White Paper No. 1). Web Science Institute. https://eprints.soton.ac.uk/428276/1/WSI_White_Paper_1.pdf

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action* (p. 280). Cambridge University Press. https://doi.org/10.1017/CBO9780511807763

Papadimitropoulos, E. (2018). Commons-based peer production in the work of yochai benkler. *TripleC*, *16*(2), 835–856. https://doi.org/10.31269/triplec.v16i2.1009

Poullet, Y. (2020). *From open data to reverse PSI – A new European policy facing GDPR* (No. 11; European Public Mosaic). Public Administraiton School of Catalonia. http://www.crid.be/pdf/public/8586.pdf

Powles, J., & Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. *Health and Technology*, *7*(4), 351–367. https://doi.org/10.1007/s12553-017-0179-1

Prainsack, B. (2019). Logged out: Ownership, exclusion and public value in the digital data and information commons. *Big Data & Society*, *6*(1). https://doi.org/10.1177/2053951719829773

Productores de música de españa (promusicae) v. Telefónica de españa SAU, C-275/06 (Court of Justice of the European Union 2008).

Purtova, N. (2017). Health Data for Common Good: Defining the Boundaries and Social Dilemmas of Data Commons. In S. Adams, N. Purtova, & R. Leenes (Eds.), *Under Observation: The Interplay Between eHealth and Surveillance* (Vol. 35, pp. 177–210). Springer International Publishing. https://doi.org/10.1007/978-3-319-48342-9_10

Purtova, N. (2021, July). *Important questions to answer before talking of 'data commons' [Written input]*. Socializing Data Value: Reflections on the State of Play [Roundtable], IT for Change, India office. https://itforchange.net/sites/default/files/2021-06/Nadezhda-Purtova-Socializing-Data-Value-Provocation.pdf

Rossi, A. (2018). How the Snowden revelations saved the EU general data protection regulation. *International Spectator*, *53*(4), 95–111. https://doi.org/10.1080/03932729.2018.1532705

Ruhaak, A. (2020, February 13). When one affects many: The case for collective consent [Essay]. *Mozilla Foundation*. https://foundation.mozilla.org/en/blog/when-one-affects-many-case-collective-consent/

Sadowski, J., Viljoen, S., & Whittaker, M. (2021). Everyone should decide how their digital data are used—Not just tech companies. *Nature*, *595*(7866), 169–171. https://doi.org/10.1038/d41586-021-01812-3

Sandoval, M. (2020). Entrepreneurial activism? Platform cooperativism between subversion and co-optation. *Critical Sociology*, *46*(6), 801–817. https://doi.org/10.1177/0896920519870577

Schiller, D. (1988). How to think about information. In V. Mosco & J. Wasko (Eds.), *The political economy of information*. University of Wisconsin Press.

Susha, I., & Gil-Garcia, J. R. (2019). A Collaborative Governance Approach to Partnerships Addressing Public Problems with Private Data. *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2892–2901. https://doi.org/10.24251/HICSS.2019.350

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, *4*(2). https://doi.org/10.1177/2053951717736335

Verhulst, S. G., Zahuranec, A. J., Young, A., & Winowatan, M. (2020). *Wanted: Data stewards. (Re-)defining the roles and responsibilities of data stewards for an age of data collaboration* [Position paper]. TheGovLab. https://thegovlab.org/static/files/publications/wanted-data-stewards.pdf

Von Hannover v. Germany (no. 2), Nos. 40660 (European Court of Human Rights 2012).

Weber, S. (2017). Data, development, and growth. *Business and Politics*, *19*(3), 397–423. https://doi.org/10.1017/bap.2017.3

Wiewiórowski, W. (2019, December). Sharing is caring? That depends… [Blog post]. *European Data Protection Supervisor*. https://edps.europa.eu/press-publications/press-news/blog/sharing-caring-depends_en

Winter, J. S., & Davidson, E. (2019). Big data governance of personal health information and challenges to contextual integrity. *Information Society*, *35*(1), 36–51. https://doi.org/10.1080/01972243.2018.1542648

Yilma, K. (2017). Digital privacy and virtues of multilateral digital constitutionalism—Preliminary thoughts. *International Journal of Law and Information Technology*, *25*(2), 115–138. https://doi.org/10.1093/ijlit/eax001

Zoboli, L. (2020). Fueling the european digital economy: A regulatory assessment of B2B data sharing. *European Business Law Review, Forthcoming*. https://doi.org/10.2139/ssrn.3521194

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, *30*, 75–89. https://doi.org/10.1057/jit.2015.5

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile Books.

Zygmuntowski, J. J. (2018). Commoning in the digital era: Platform cooperativism as a counter to cognitive capitalism. *Praktyka Teoretyczna Numer*, *1*(27). https://doi.org/10.14746/prt.2018.1.7

Zygmuntowski, J. J. (2020). *Kapitalizm sieci*. Roz:Ruch Publisher.

Published by


in cooperation with