A fourth way to the digital transformation: The data republic as a fair data ecosystem

Calzati, S.; van Loenen, B.

DOI
10.1017/dap.2023.18

Publication date
2023

Document Version
Final published version

Published in
Data & Policy

Data & Policy (2023), 5: e21
doi:10.1017/dap.2023.18

RESEARCH ARTICLE

A fourth way to the digital transformation: The data republic as a fair data ecosystem

Stefano Calzati and Bastiaan van Loenen

Urban Data Science Group, Department of Urbanism, TU Delft, Delft, the Netherlands
Corresponding author: Stefano Calzati; Email: s.calzati@tudelft.nl

Received: 05 December 2022; **Revised:** 25 April 2023; **Accepted:** 23 May 2023

Keywords: Data commons; data governance; ecosystem; EU; fairness

Abstract

To harness the promises of digital transformation, different players take different paths. Departing from corporate-driven (e.g., the United States) and state-led (e.g., China) approaches, in various documents, the European Union states its goal to establish a citizen-centric data ecosystem. However, it remains contentious the extent to which the envisioned digital single market can enable the creation of public value and empower citizens. As an alternative, in this article, we argue in favor of a *fair* data ecosystem, defined as an approach capable of representing and keep in balance the data interests of all actors, while maintain a collective outlook. We build such ecosystem around data commons—as a third path to market and state approaches to the managing of resources—coupled with open data (OD) frameworks and spatial data infrastructures (SDIs). Indeed, based on literature, we claim that these three regimes complement each other, with OD and SDIs supplying infrastructures and institutionalization to data commons' limited replicability and scalability. This creates the preconditions for designing the main roles, rules, and mechanisms of a data republic, as a possible enactment of a fair data ecosystem. While outlining here its main traits, the testing of the data republic model is open for further research.

Policy Significance Statement

Today's European Union digital strategy over-emphasizes the defense of individual rights, overlooking the societal-level impact of data. It is necessary to calibrate data governance at local and (supra)national levels striking a balance between individual and collective rights, principles, and values. To preserve democracy and the rule of law before digital transformation, data republicanism is a viable path as it devises checks and balances to account for the distribution of data power. This implies to design mechanisms to connect local and (supra) national data actors and practices, as well as to establish the processes of data stewardship and data arbitration able, the former, to build technological and legal capabilities in the public sector as well as data literacies in the citizenry, and, the latter, to adjudicate contentious cases whenever tensions among different data interests arise.

1. Introduction

Data have become the key enabler of the fourth industrial revolution (Jasperneite, 2012). In fact, data are considered as a pivotal asset of the digital transformation which impacts in all sectors, boosting innovation (European Commission, 2013) and creating economic value (European Commission, 2014). However, *how* such innovation and value can be best achieved is a matter of data governance (Micheli et al., 2020) and subject to different approaches.

At present, it is possible to identify three major visions when it comes to governing digital transformation (Calzada and Amarall, 2020; Schneider, 2020; Calzati, 2023): (a) a corporate-driven approach (e.g., United States) based on market deregulation and favoring economic competitiveness among tech stakeholders and platforms (Turner, 2017; Martin et al., 2018); (b) a state-led approach (e.g., China) depending on authority-defined plans and striving for global techno-economic leadership in strategic sectors while maintaining state control over social and moral behaviors (Au and Kuuskemaa, 2019; Roberts et al., 2021); (c) a citizen-centric approach (e.g., European Union [EU]) aiming to achieve digital transformation by safeguarding human rights and balancing economic competitiveness with social inclusiveness, democratic participation, and environmental sustainability (von der Leyen, 2019; European Commission, 2022).

While the corporate-driven and state-led approaches have shown limitations concerning, in the former case, the lack of contextual adaptability of technology (Kummitha, 2020) and limited social inclusiveness (Kingston and Viitanen, 2014; Sanfilippo and Frischmann, 2023) and, in the latter case, the stiffing of technological innovation by encumbering bureaucracy (Sun, 2007; Fu et al., 2016) and the hindering of R&D diversification due to lack of incentives (Han et al., 2019; Genin et al., 2021), the European path too faces shortcomings. On the one hand, the envisioned digital single market (European Commission, 2021) puts *private actors* at the center of the stage, questioning both how the creation of public value can be accounted (Taylor, 2021), as well as the role of citizens and democratic participation in this scenario (Cardullo and Kitchin, 2019). On the other hand, the focus on citizens as individuals risks overlooking the impact of data at collective level (Smuha, 2021; Viljoen, 2021), leaving unsolved the issue of how to comply with values, such as democratic participation or environmental sustainability, that pertain to society in its entirety.

This is why these tensions within the European context demand a paradigmatic shift in the way to govern data (Micheli et al., 2020). In this regard, scholars have called for the design of a *comprehensive* framework (Kozminski et al., 2021) which moves away from an understanding of data governance as either targeting certain actors over others—for example, citizens, public actors, and private actors—or prioritizing one value over others—oftentimes economic competitiveness over social inclusiveness or environmental sustainability. To do so, such comprehensive framework must first consider the data landscape as an ecosystem (van Loenen et al., 2021) that, by definition, is irreducible to any of its actors or values for its sustainable working; and secondly design strategies to keep the whole ecosystem in balance by redressing possible power asymmetries arising among actors or conflicts among values. At stake is the reconsideration of data governance from an actor-network approach (cf. Latour, 2004) to a systemic-procedural one.

From these premises, in this article, we elaborate the idea of a *fair* data ecosystem as a governance model where the "data interests" (Hasselbalch, 2021) of all actors are *systemically* taken into account and disentangled based on rules and mechanisms that adjudicate which values and actors are to be prioritized on a case-by-case basis—indeed, a data republic. To operationalize the data republic, we propose to couple a data commons (DC) approach with Open data (OD) frameworks and spatial data infrastructures (SDIs). On the one hand, DC is regarded as a viable third path to market and/or state approaches to the managing of data (Morozov and Bria, 2018; Bangratz and Förster, 2021), with the intent to reappropriate data by citizens and repurpose these data by keeping a societal outlook in sight. So far, however, DC initiatives remain affected by limited replicability (de Lange and de Waal, 2019) and scalability (Calzada and Almirall, 2020). On the other hand, OD and SDIs initiatives have consolidated over the last three decades backed up at both institutional and infrastructural levels (Welle Donker and van Loenen, 2017; Mulder et al., 2020; Raymond and Kouper, 2023). And yet, these initiatives miss the needed context-flexibility to respond to locals' data needs and involve them in the provision of indigenous data (Lupi et al., 2020; Verhulst et al., 2020; Valli Buttow and Weerts, 2022). Hence, taken together, these three regimes provide mutually complementary features, enabling "institutioning" (Huybrechts et al., 2017) and "infrastructuring" (Ludwig et al., 2018) processes, intended as the dynamic interplay between top-down and bottom-up stances and values. This creates the preconditions for designing the main roles, rules, and mechanisms of the data republic as a fair data ecosystem. At this stage, the envisioned data republic

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

remains at a high level of abstraction, whose testing will be the focus of further research. In other words, the data republic outlined here is a theoretical setup: we do not claim that the model as such will fix data governance problems; instead, we advance that, based on the limitations of current approaches to digital transformation (the United States, China, and the EU) and regimes (OD, SDIs, and DC), the data republic constitutes a *possible* setup to overcome these limitations.

The article is organized as follows: In Section 2, we outline the three approaches to digital transformation—corporate-driven, state-led, and citizen-centric—discussing their goals and main limits. In Section 3, we define systemic fairness in the context of data ecosystems. Section 4 reviews OD, SDIs, and DC, highlighting strengths and weaknesses of each regime and Section 5 advances the coupling of these regimes based on their complementarity. In Section 6, we operationalize such coupling by detailing the main features of the data republic. Section 7 provides concluding remarks and outlines research ahead.

2. Three Approaches to Digital Transformation

When it comes to the governing of digital transformation, it is possible to identify three major competing visions. Schneider (2020) notes that, while the United States and China are often regarded as the two dominant governance models, leading some scholars to advance the idea of a new cold war (Lippert and Perthes, 2020), the EU has strived to carve for itself a "third way" (Bendiek and Schallbruch, 2019), intended as a form of digital self-determination. At stake is not only a matter of technological advancement, but also, more radically, of economic and geopolitical supremacy (Pohle and Thiel, 2020; Voss and Pernot-Leplay, 2023). To achieve this, major actors contribute to shape the digital transformation across scales and in different socio-political contexts (Calzati, 2021), often negotiating between fundamental rights (e.g., in the EU), economic competitiveness (e.g., in the United States) and societal "harmony" (e.g., in China).[1]

2.1. The United States corporate-driven approach

Calzada and Amarall (2020) write that in the United States the so-called GAFA (Google, Amazon, Facebook—now Meta—and Apple) are the symbol of a paradigm "driven by large technological private multinationals who collect massive amounts of data from global citizens without any informed consent." This is a *corporate-driven* approach to digital transformation in which the public sector tends to either play a facilitating role or become de facto the client/recipient of tech solutions developed, controlled, and owned by private corporations. In this scenario, the blossoming of big tech corporations "is seen as positive both for innovation and economic growth and hence is fostered," leading to "extremely high revenues [that] allow these companies also the power to lobby governments" (Schneider, 2020). While favoring a competitive landscape where innovation and economic success go hand in hand, this approach shows drawbacks especially due to the lack of contextual adaptability of the developed technologies, as well as their limited social inclusiveness.

Concerning the first kind of drawbacks, Kummitha (2020) notes that because "corporate firms (…) sell the very same technologies they developed for different cities (…) technologies often ignore place-based differences and the local cultural and community context." The lack of attention to contingent needs produces discrepancies along the socio-technical axis which is then up to the public sector to tackle, if it has the capacity to do so. Kalpokas (2022) warns, in this regard, that the enduring gap between the technical and the social, if not adequately tamed, produces increasing inequalities and forms of discrimination. This links to the second kind of drawbacks, notably the limited social inclusiveness following the implementation of data-driven technologies. On this point, Viitanen and Kingston (2014) argue that "inequality and poverty do not often feature in debates, but the technological fixes will have distributional

[1] From a normative standpoint, here, the focus is on the United States, China, and the EU. For a discussion that unpacks the techno-geopolitical tensions between big tech corporations and states across a transnational axis, see Wen (2021) and Glasze et al. (2023).

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

consequences under which there are winners and losers." When put in context, a corporate-driven approach risks overlooking group-level unwanted social effects of digital transformation, especially toward socio-economically disadvantaged people. As a Pew report (Pew Research Center, 2020) puts it, technology tends to further empower the already powerful and to further "diminish" those who are already disempowered. This is so especially in the United States where the corporate-driven approach inevitably keeps economic competitiveness and profit as main principles, not only granting to private actors wide and discretional power over which (social) issues to address and which (technical) solutions to deliver, but, more problematically, "without considering the full range of consequences" (Sanfilippo and Frischmann, 2023) of their own innovations.

2.2. China's state-led approach

By contrast, China's *state-led* approach is regarded as technological nationalism (Jiang and Fu, 2018) and is meant to foster harmony and social stability (Au and Kuuskemaa, 2019). Within this approach, public authorities create new lanes for digital transformation via a top-down logic in which it is up to state authorities to broadly dictate the direction to follow. Heavy public backing through financial facilitations favors the achievement of mid- to long-term targets, all part of China's goal to reach technological market supremacy by 2025 (Hausstein and Zheng, 2018). Yet, this approach too presents shortcomings. Bureaucracy tends to stiff innovation due, on the one hand, to the enduring "fragmentation of the state governance structure [and] the poor coordination within the bureaucracy" (Sun, 2007) and, on the other hand, to a bottleneck disadvantaging innovation by small and medium enterprises compared to big state-owned firms (Fu et al., 2016). Moreover, despite having transited from a manufacturing "copycat" model to an indigenous technological paradigm (Yu et al., 2017; Lee, 2019), China's state-funded digital transformation still misses to foster a strong link between industry and research because of lack of incentives to experiment out of identified paths (Han et al., 2019). Specifically, Genin et al. (2021) show that state-owned tech enterprises are more unlikely to adapt to market competition and experiment with innovation than other firms. This—scholars (Yu et al., 2017) contend—has repercussions on enduring gaps, in terms of capacity and willingness to innovate, between China and other international players such as Japan, Germany, and the United States. In this respect, Zeng (2017, p. 70) calls for "a legal and regulatory system that encourages (…) open and fair competition among private, state-owned, and foreign enterprises." Currently, however, China's state-led vision on digital transformation maintains public authorities in a near-monopolistic position concerning how to steer such process beyond or even against the interests of the various players involved.

2.3. The European Union citizen-centric approach

In the Declaration on Digital Rights and Principles (DDRP), the European Commission (2022) reaffirms the objective to "promote a European way for the digital transition, putting people at the center. It shall be based on European values and benefiting all individuals and businesses." It is certainly not the first time that the EU claims to pursue a *citizen-centric* approach to digital transformation. Since 2014, the EU has taken steps in this direction, with initiatives such as the Cybersecurity Act (European Parliament and Council, 2013b), the General Data Protection Regulation (European Parliament and Council, 2016), the Regulation on the Free Flow of Non-personal Data (European Parliament and Council, 2018), the Ethics Guidelines on Trustworthy AI (High-Level Expert Group on Artificial Intelligence, 2019), and the Data Governance Act (European Parliament and Council, 2020b). These steps are part of a digital strategy that aims to keep the EU abreast of competitors (the United States and China) while preserving its core values, as pinned down in the DDRP: (1) preserve people's rights; (2) support solidarity and inclusion; (3) ensure freedom of choice; (4) foster democratic participation; (5) increase safety, security, and empowerment of individuals; and (6) promote sustainability. Noteworthy is that these values equally split between a half (1, 3, 5) focusing on the individual and the other half (2, 4, 6) pertaining to the society. Hence, the DDRP does strive to strike a balance between subject-centric and collective-centric dimensions.

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

Based on these premises, the EU approach presents criticalities when turned into practice. On the one hand, recently released documents question the extent to which the EU is really pursuing a citizen-centric approach; on the other hand, it is the individualistic connotation of such approach to come under scrutiny, due to its inability to capture the societal dimension of digital transformation and to comply with values that applies to society as a whole.

For instance, the European strategy for data (European Parliament and Council, 2020a) is linked to the aim of "creating a single market for data that will ensure Europe's global competitiveness and data sovereignty." To speak of a "market" underpins the idea of data as a commodity, which is a very contentious idea to begin with for two reasons. On the one hand, it considers by default data as something to be seized, owned, and exchanged under an economic and proprietary rationale. On the other hand, the idea of market smoothly turns subjects—either physical or legal—into consumers. This has profound implications on the concrete enactment of those European values—for example, social inclusiveness, environmental sustainability, and democratic participation—which cannot be boiled down to an economic framing. Along the same line, in the 2021 Digital Europe Programme (European Commission, 2021), a "data ecosystem" is defined as "a platform that combines data from numerous providers and builds value through the usage of processed data." Discussions on "platformization" (van Dijck et al., 2019; Cristofari and Helmond, 2023) have widely unveiled the commodification of services and actors that such concept puts forth. In fact, the EU considers platformization as a default phenomenon, allowing to re-enact in the digital realm forms of private monopoly and abuse of market power. Not surprisingly, the same document specifies the goal of "building a thriving ecosystem of private actors to generate economic and societal value from data" (European Commission, 2021). This approach aligns more closely with a corporate-driven one, which —as discussed above—has been shown to present limitations. On this point, Taylor (2021) warns against the notorious difficulty of "establishing meaningful accountability for the private sector" which de facto hinders an effective public scrutiny of how big tech giants operates, for which purposes, and with which results. Hence, a reconsideration of what citizen-centric really means for the EU becomes necessary, insofar as the way in which such approach is currently enacted raises concerns over the compliance with the European principles enshrined in the DDRP. In fact, it is safe to say that since its inception the EU's data governance has been centered around the boosting of economic value and the preservation of individuals and their rights as consumers (Valli Buttow and Weerts, 2022). In this context, citizenship has by and large been coopted as a glamoring rather than a pivotal concept.

To purse a citizen-centric approach, on the one hand, it is necessary to move beyond a chiefly economic rationale. This can be achieved by rethinking democratic participation *through* and *about* digital transformation. As Cardullo and Kitchin (2019) note, we need to redesign participation toward "more extensive public consultation, collaboration and co-production" which are rooted in "a set of civil, social, political, symbolic and digital rights and entitlements," rather than in market principles.

On the other hand, it is crucial to inscribe the EU's right-based standpoint on data governance (Donahoe and Metzger, 2019)—that is, a standpoint that chiefly protects individual rights, such as privacy and freedom of choice—within a broader perspective that accounts for the collective-level dimension of digital transformation and related societal values, that is, a dimension that cannot be reduced to the sum of individuals and their rights. While a right-based standpoint might constitute the *necessary* baseline to individual autonomy, there is increasing evidence that this approach cannot be exhaustive (Smuha, 2021; Viljoen, 2021). For instance, Viljoen (2021) notes that the individualistic vision behind the current EU approach does not account for the relational nature of data and the consequent trade-off effects that data reuse involving two subjects might have on unaware third parties. On this wave, Smuha (2021) suggests taking inspiration from environmental law for tackling potential collective-level effects caused by digital transformation, such as the erosion of the functioning of the rule of law, which can be neither accounted for nor mitigated by current individualistic approaches to digital transformation. Hence, while a human right-based approach to digital transformation is fundamental to protect the individual's autonomy, it is *insufficient* to protect Europeans as a whole.

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

3. Characterizing a Data Ecosystem and Its Fairness

To the extent to which digital transformation is a process increasingly affecting all fields of life, a proper governance framework that is meant to regulate such process must abandon a targeted approach to either actors or values and adopt an ecosystemic standpoint instead.

An ecosystem is characterized by interacting biotic and non-biotic elements, so that its behavior cannot be studied by isolating either elements or interactions; rather, it must be studied in its entirety. Similarly, a data ecosystem is a concept framing the sociotechnical elements, actors and procedures contributing, all together, to create and manage data-based initiatives (Jarke et al., 2019). As for fairness, a general definition from the *Cambridge Dictionary* characterizes it as "the quality of treating people equally or in a way that is right or reasonable [by] considering everything that has an effect on a situation, so that a fair judgment can be made." From these premises, a fair data ecosystem can be understood as one able to balance out the data interests (Hasselbalch, 2021) of all the actors in play, based on shared values and in view of the sustainability of the whole data landscape (van Loenen et al., 2021). This is an ecosystemic characterization of fairness that underscores the trading off among different interests to seek to maintain an overall equilibrium.

This understanding of fairness moves away from both a reductionist and an essentialist definition of the term. Within the first group fall all those attempts, especially in computer science and software engineering, which seek to provide a mathematical definition of fairness, based on metrics against which algorithms can be tested on a pass/fail basis. Notwithstanding, the plethora of mutually exclusive characterizations of fairness this approach leads to (Kleinberg et al., 2016), its main limitation rests in the reduction of fairness to a computational matter which overlooks the contextual dimension of the term (Lee and Singh, 2021). Far from being a purely quantitative affair, an evaluation of fairness cannot prescind from taking into account contingent factors, from individual to cultural values.

On the other hand, an essentialist understanding of fairness does account for the context-dependency of digital transformation, and yet it remains trapped within an understanding of fairness as a core quality of a technology or data process, so that it can be objectively assessed. For instance, Lee et al. (2021) characterize fairness as an "evaluative judgement of whether a decision is morally right," from which they proceed to identify key ethical indicators for enabling a "customized measurement of what 'fair' looks like" in each context. In so doing, however, fairness keeps attached to a specific scenario, falling short of producing a comprehensive approach to digital transformation.

To shift toward an ecosystemic understanding of fairness it is worth looking at how the EU defines this term in the context of the development and implementation of data-driven technologies. Notably, the High-Level Expert Group on Artificial Intelligence (2019) disentangles fairness as both a substantive and procedural affair: on the one hand, "individuals and groups are free from unfair bias"; on the other hand, it must be guaranteed "the ability to contest and seek effective redress" against tech-based decisional processes. This double-sided understanding implies not only the recognition of one's own diversity and their inclusion by default into the collective, but the enforcing of mechanisms to reclaim such individual agency as part of a whole. On this point, Rochel (2021) notes that as a structuring principle of the GDPR "fairness" is "linked to principles such as proportionality and other procedural dimensions of a balancing exercise involving rights and interests." This highlights well the fact that, beyond the matching of certain requirements, fairness is an act of balance based upon the recognition and negotiation among different interests and rights on a flexible and rolling basis. A fair data ecosystem, then, shall be regarded not much as an arena where different players are connected, but as a process that constantly reshapes its own power relations.

A governance framework that aims to regulate a data ecosystem *fairly* identifies roles and rules to *systemically* represent the data interests of all actors involved, as well as mechanisms to adjudicate situations where conflicts among actors and/or values might arise. Hence, we introduce the idea of data republic as a concretization of such fair data ecosystem. "To be a republican," Susskind (2022) notes, "is to regard the central problem of politics as the concentration of unaccountable power." Therefore, a data

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

republic exists to the extent to which roles, rules, and mechanisms are envisioned to keep the whole ecosystem in balance, by adjudicating priorities in contentious cases.

In the following section, we will make a case for the coupling of a DC approach to the governing of data with OD frameworks and SDIs as one promising way to enable a fair data ecosystem.

4. Data Commons, Open Data, and Spatial Data Infrastructures

4.1. Data commons: Prospects and challenges

As a regime for the managing of resources, the concept of the commons can prove fruitful in the context of data governance.

Originally, the "commons" referred to common-pool resources (CPRs)—such as fisheries or forests—which are characterized by non-excludability and rivalry. These terms point to the fact that: (a) it is difficult to forbid access and use of CPRs to any potential beneficiary; and (b) the use of CPRs depletes them and reduces further use by others. Ostrom (1990) showed that the self-management of CPRs by communities can be more efficient and sustainable than market-driven or state-led approaches, provided that formal and informal principles, roles, and rules are designed and abided to.

By now, the commons has spilled over onto realms other than CPRs, coming to identify more broadly a system consisting of a resource, its users, the institutions binding them, and the associated mechanism processes (Feinberg et al., 2021). The trading mark of such system is to be non-appropriative by default (knowledge, technology, assets, and outputs are not owned, in the commercial sense of the term, but summoned up and recirculated); collaborative by design (it considers all actors and links within the ecosystem as integral and necessary to the system's flourishing), and collectively sustainable in its goals (indeed, commons for the community). In other words, commoning practices enact an ecosystemic approach that strives to balance out individual and collective interests and values.

When it comes to data and technology, the spillover of the concept of the commons (Shkabatur, 2019) was favored by the consolidation of the Internet—an open infrastructure—supplying the basis for the proliferation of new forms of co-innovation, via freely accessible knowledge, design, and software (Kostakis et al., 2015). DC characterizes a regime in which actors join forces in the collection, pooling, and use of data (and digital infrastructures) subservient to the delivery of services for the whole community. DC initiatives (Morozov and Bria, 2018) aim to counteract and/or repurpose the centralized ownership and use of data—either by tech companies or states—by giving these back to citizens, with the goal to foster sustainable collective data practices. DC initiatives, then, truly reinserts citizens into the data ecosystem and allow them to have a say on or also co-develop tech solutions. At present, however, these initiatives fail to consolidate and achieve replicability (de Lange and de Waal, 2019) and scalability (Calzada and Almirall, 2020).

Over the last few years, various initiatives (Balestrini et al., 2017; Ajuntament de Barcelona, 2019; de Lange and de Waal, 2019) have emerged around Europe attempting to enact a commons-inspired approach to data (and digital infrastructures) usually at city scale. For instance, Wolff et al. (2019) explored ways of creating more awareness in Milton Keynes's citizenry about what can be done with and through data. Their research shows that digital platforms can help urban communities gather around shared concerns and proactively advance solutions. However, data literacy is still limited in the population, requiring initiatives to tackle such scarcity through the institution of facilitating roles for connecting governing bodies with communities. Mulder and Kun (2019), instead, investigated the extent to which the pooling of communities via co-creative partnerships can lead to sustainable initiatives that integrate and/or rework institutional processes in the long run. They show that these initiatives are largely effective to boost collaboration at local level and on a temporary basis, but they fail to put forth a systemic change. Similarly, the "Bristol approach" (Balestrini et al., 2017) takes a participatory stance, prefiguring ways-of-doing which put citizens' needs at the center of data-driven solutions. While enabling co-design, this approach falls short of identifying a proper governance framework which can sustain or replicate these initiatives. Overall, while much effort goes into the repurposing of data and infrastructures, this

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

effort is still hardly reconciled with the fluid, self-organizing drives of communities (de Lange and de Waal, 2019).

Institutionally speaking, one of the most consistent examples from the perspective of a fair data ecosystem comes from Barcelona (Ajuntament de Barcelona, 2019). In 2016, the Catalan municipality launched a "new social pact on data" composed of various initiatives, among which the establishment of DC regimes allowing citizens to own and keep control over their data. In the words of Morozov and Bria (2018), the goal was to make good use of the power of data through "an ethical and responsible innovation strategy, preserving citizens' fundamental rights and information self-determination." Despite the echo produced by Barcelona's initiative, several barriers are still in place (Monge et al., 2022), most of which resonate with those identified above, especially the difficulty of implementing long-term sustainability due to the lack of institutionalization and the needed infrastructures. Overall, Barcelona's case reveals that a commons approach can be fruitfully applied to data only to the extent a broader ecosystem is taken into account.

4.2. Open data

To achieve the economic benefits of digital transformation, OD constitute a crucial asset (European Commission, 2014). Data are considered open when they are not personal and they can be freely used, re-used, and re-distributed by anyone, at most restricted by the obligation to name sources and "share-alike" (Open Knowledge Foundation, 2013). After over a decade of political and technological pooling, the push toward OD has acquired an institutional status in many countries.

In 2003 (and in an updated form in 2013), the European Parliament and Council (2003, 2013a) released the directive on the reuse of public sector information (PSI directive), with the main goal to boost economic value through the reuse of such information. Subsequently, in 2019, the EU enacted the OD directive (European Parliament and Council, 2019), which enlarges the scope of the PSI directive to involve, for instance, research data, as well as identifies priority sector data to be released as OD (e.g., geospatial, statistics, and mobility). Recently, the Data Governance Act (European Parliament and Council, 2020b) and the Data Act (European Parliament and Council, 2022) were proposed as policy pillars to further boost data sharing, fostering a trustworthy European digital single market. In fact, the Data Governance Act establishes (a) measures to facilitate the (re)use of sensitive public sector data; (b) mechanisms for citizens and businesses to make their data available; and (c) cross-border and cross-sector data sharing. The Data Act specifies the actual rights on the access to and (re)use of data, notably granting users to gain access to and make use of data generated by them, as well as identifying avenues for public sector bodies to access and use private sector data in exceptional circumstances (such as a public emergency). While these directives signal an increasing drive toward OD and the fostering of a data-inclusive ecosystem in terms of both actors involved and types of data pooled, limitations remain not only concerning the ethos surrounding such directives (as discussed in Section 2) but also their implementation.

Retracing the evolution of OD directives, Verhulst et al. (2020) identify three waves. The first wave consisted mainly of making national data available upon request to an audience of experts, such as lawyers, journalists, and researchers. In the second wave, open government data (OGD) were made open by default to anyone. Yet, aware that data supply alone does not lead to more (re)use, nor by default to the creation of public value, the (expected) third wave embraces a more purpose-driven approach, putting equal emphasis on data supply and the broader context in which data are meant to be (re)used. Indeed, as Welle Donker and van Loenen (2017) stress, it is important to be in touch with societal issues, while matching demand and supply of data. According to the authors, preconditions are for data to be: (a) known to the user; (b) attainable by the user; and (c) usable for the intended purpose of the user. Clearly, at stake is a matter of knowing which data are needed and for which (re)use purposes. On this point, Lupi et al. (2020) note that we need "appropriate data" rather than simply OD, insofar as today's enduring under-exploitation of OD seems mainly due to a misalignment between "the provision of data and the actual information needs of local actors."

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

To this, it must be added that OD initiatives have so far chiefly focused on the national and supranational levels, while much data reside at local level (Verhulst et al., 2020). This is also why scholars have called to actions to mobilize authorities at various levels for not only making datasets available, but also engaging citizens and foster stakeholder communities around OD (Mergel et al., 2018). Hence, while OD represent a key and by now institutionalized enabler to boost digital transformation, what is still missing are mechanisms and practices connecting (supra)nation and local levels favoring the matching between provision and demand of data across scales.

4.3. Spatial data infrastructures (SDIs)

A SDI is a dynamic and multi-disciplinary architecture that allows for access, reuse, and sharing of spatial data (Crompvoets et al., 2008). Originally, SDIs consolidated following the proliferation of geospatial information systems which allowed governments for an increasing amount of spatial data to be collected for analytical and policy purposes (Scott and Rajabifard, 2017). SDIs, then, tend to have a national dimension by design with public authorities responsible for coordinating ready access to and interoperable use of these data.

As complex architectures, SDIs consists of seven core elements (van Loenen, 2006): (1) (georeferenced) data; (2) people (actors who collect, create, process, and/or use data); (3) policies (for the allocation, use, and circulation of data); (4) institutional frameworks (political support and institutional arrangements to enact the use of data and pool actors into collaboration); (5) technology (methods and instruments, for the collection, gathering, storing, use, and circulation of data); (6) standards (the specifications, quality, and requirements for a smooth circulation of data and the interoperability across different services and actors); and (7) financial resources (the systematic allocation of fundings to keep SDI operations going). Today, by leveraging on over three decades of development, SDIs represent a robust backbone for (supra)national data-driven initiatives.

The three waves of OD echo the phases identified by Vancauwenberghe et al. (2018) concerning SDIs' evolution. In the 1990s, SDIs were producer-driven, with the focus exclusively on the supply of public georeferenced data (Masser, 1999) by national bodies; in the early 2000s, some SDIs have become process-driven transiting from the provision of data to the provision of web-based services (Budhathoki et al., 2008); today, a user-centric approach is advocated for in the attempt to not only match users' needs with the right kind of data (and how these data can be best delivered), but also to involve users in the collection and release of data.

It is especially in this latest regard that OD play a crucial role, leading to the idea of open SDIs (SPIDER Consortium, 2021), whereby "open" refers as much to data from the public sector, private sector, and citizens (Vancauwenberghe and van Loenen, 2018), as to the infrastructure itself (Vancauwenberghe et al., 2018). Hence, an open SDI can be conceptualized as a data ecosystem where "all stakeholders commonly govern, share, and use open geodata" (SPIDER Consortium, 2021). While at the core of open SDIs are spatial data, in principle, the concept can be extended to other data—such as energy data or health data—provided that they are georeferenced. Welle Donker et al. (2016) show, for instance, that energy data belonging to a private energy distributor endowed with a public task were successfully released in an open manner while preserving users' anonymity by setting different scales of granularity. Beyond such pilot improvements, however, at present issues remain concerning how to concretely operationalize such openness of SDIs, especially with regard to the systemic attuning to and harnessing of the local scale.

Pivotal to the consolidation of SDIs in the EU is the Infrastructure for Spatial Information in the European Community (INSPIRE) directive (European Parliament and Council, 2007) which aims to establish shared requirements for European SDIs in view of environment-related policies and activities. The directive sets standards and duties for public bodies to produce, receive, manage, or update spatial datasets, with the goal of facilitating the access to and sharing of spatial data across Europe and favor cross-boundary policymaking. In this respect, INSPIRE is a model of replicability—insofar as it is implemented by each member state—as well as scalability, insofar as its data-related specifications apply at all levels. However, 15 years after the approval of the directive, the number of users making use of

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

INSPIRE-compliant spatial data for public tasks is still marginal, raising questions over the possible mismatch between the product (i.e., data) and the process (i.e., their use). In this regard, combining open SDIs with a DC approach that puts by default people and the community at the center of data practices constitute a matching worth exploring.

5. Coupling Open Data and Spatial Data Infrastructures with the Commons

Table 1 summaries strengths and weakness of DC, OD, and SDIs, and related values (keeping the DDRP as a benchmark). Their complementarity is key to fostering a fair data ecosystem. On the one hand, literature underscores how DC initiatives put communities at the center of their practices, showing a constitutive granularity which gives agency to local issues and actors. As a regime for managing data, the commons embodies an ecosystemic approach that tries to balance out the data interests of the community, in view of common(ing) goals. At the same time, however, these initiatives present limited replicability and scalability, because, being confined to a micro-scale dimension, they have limited institutionalization and infrastructures.

On the other hand, an increasing amount of OD has been released since late 2000s, backed up by a series of directives, policy documents, and organs. This institutional framing mirrors the development of SDIs, which have consolidated, since early 1990s, as a backbone for the collection and dissemination of georeferenced data. At present, however, open SDIs remains at a conceptual level of development still seeking an operational approach capable of attuning to context-based demand and use of data, as well as linking (supra)national policies and infrastructures to the local dimension.

Here, we argue that the coupling of these three regimes does enable their mutual "institutioning" (Huybrechts et al., 2017) and "infrastructuring" (Ludwig et al., 2018), intended as dynamic processes merging top-down and bottom-up stances and values. This is what makes sure to foster a fair data ecosystem. In other words, to think ecosystemically implies to establish the conditions—for example, links between institutional and noninstitutional initiatives; data literacies and techno-legal capacities; trust across actors—to turn data governance into a systemically fair process.

Far from being "monolithic" organizations, institutions are highly dynamic, with change being continuously exerted from both the outside and inside (Streeck and Thelen, 2005). In this respect, Huybrechts et al. (2017) coin the gerund "institutioning" for accounting for such dynamic interplay and the negotiation of competing (non)institutional stances. Indeed, research (Cazacu et al., 2020) shows

Table 1. *Strengths and weaknesses of open data, spatial data infrastructures, and data commons and related values*

	Pros/promoted values	Cons/missing values	
OD	Institutionalization at (supra)national level; economic value	No focus on contextual purposefulness of data and local communities' needs; limited participation	OPEN SDI FAIR DATA ECOSYSTEM
SDIs	Infrastructural robustness at (supra)national level; economic value	Lacking mechanisms to give agency to and harness the local scale; limited participation	
DC	Pooling of local communities' expertise and needs; socio-economic and participatory values	Lack of scalability (beyond the local dimension) and systematic replicability (in different contexts)	

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

that the process of institutionalization (e.g., of new ideas, expertise, instances, and roles) occurs at best when top-down and bottom-up dynamics not only are juxtaposed, but enter in dialogue at various scales, allowing to decrease power distances and guarantee more agency to all actors in play.

In turn, the process of institutionalization is based upon and brings with itself infrastructures. An infrastructure can be considered as the basic system for enabling certain operations. Research (Star and Ruhleder, 1996) has unveiled the sociotechnical nature of infrastructures, as backbones comprising of both hardware and software arrangements—from cables, networks, and devices to protocols, services, and expertise—which allows for the reiteration of shared practices. In this case too, the noun can metamorphose into a verb, whereby "infrastructuring" (Ludwig et al., 2018) comes to characterize people's active (re)design of hardware and software components fostering, de facto, an ongoing co-development of the social *and* technological sides of infrastructures.

Having said this, if on a theoretical level OD, SDIs, and DC are complementary, on a practical level their coupling occurs at the nexus where the local–global axis and the individual-collective axis intersect. This demands forms of "collectual" governance (Calzati, 2023) that synthesize subject-centric and collective-centric tensions across various scales. To make sure that OD, SDIs, and DC can constitute the pillars of a fair data ecosystem, it is necessary to contextualize their coupling, as we will do in the section unpacking the concept of "community" and "general interest."

6. Toward the Data Republic

6.1. Community and (general) interest(s)

Beckwith et al. (2019) consider OD and the commons as two mutually exclusive regimes, arguing that "data are 'about' locals" and discussing how "making data available as OD would lead to community impacts that were most unwelcome," to the extent to which the very fact of opening data might have (un) intended consequences for both who is given access to and who remains excluded. On a closer inspection, however, such position rests on a thin conceptual basis. When speaking of (open) data and communities—and more generally of data ecosystem—one should always ask: Who is in play? According to which rules and values? Over which timeframe? At stake is an issue not of ownership, but control: the latter is open—much more than the former—to modulation. As Hummels et al. (2021) note "in the end, mitigation mechanisms are necessary for both those who incur damages due to their inclusion, and those who incur damages from being excluded."

This leads to discuss two key terms at the basis of a fair data ecosystem: "community" and "general interest." A (data) community is a fractal concept (Tannier and Thomas, 2013) as far as its scale is concerned in that it depends on the interplay among three components: infrastructures (e.g., ICTs), law (e.g., national policies, regional directives, and city's orders), and locals' knowledge (e.g., people's practices and relations relevant to and framed within a given place). As long as these three components are ideally co-extensive (i.e., they overlap), then authority and territoriality are fully legitimate, and the exercise of power coincides with (and can be scrutinized in) the interest of the whole community. Whenever the co-extensiveness of the three is not guaranteed, as it often happens—for example, a community's infrastructure extends well beyond the human relations bound to the territory or an international actor comes in play in a small community under international market laws—then we have a weakening of legitimacy because of a discrepancy between authority (who takes the decision) and territoriality (reduced or no community's agency). Here is when self-organization fades, being substituted by top-down-only or global-market approaches.

This, in turn, implies that the general interest of a community is inevitably subjected to ongoing (re) negotiation. Already today, national, and supranational legal frameworks are in place for disentangling individual and collective interests concerning the access and (re)use of (personal) data. This is so because "general interest" is an entangled concept that demands constant contextualization. From an empirical perspective, the concept reflects the diversity of interests of all actors involved in a given situation (Healey, 1997); from an ethical perspective, it constitutes the synthesis (*not necessarily* the sum) of all actors' interests (Innes and Booher, 2015). In fact, such synthesis is never given once and for all; it is based

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

on discontinuities across the community. Concretely, this demands the design of a participatory approach able to identify, negotiate, and adjudicate among such discontinuities. But what *kind* of participation? And involving *who*? According to Arnstein (2000), it is only when citizens get effective and direct accountability and deliberative powers over the decisions to be taken that participation is valuable. To have successful participation, then, it is crucial to "manage the system as a process of continuous innovation, learning and adaptation" (Toots, 2019), whereby new competences and skills are constantly acquired and put to use.

As far as the "who" is concerned, the quadruple helix—public sector, private sector, academia, and citizens—shall be regarded as the baseline instead of the optimum. In fact, a whole galaxy of (non) institutional actors does contribute to inform any data ecosystem: NGOs, nonprofit organizations, data intermediaries, data stewards, etc. (*including* free riders). This heterogeneous galaxy is increasingly acknowledged—yet, not operationalized—by the EU. For instance, the Data Governance Act (European Parliament and Council, 2020b) specifies the need to "designate one or more competent bodies to support the public sector bodies which grant access to the re-use of the categories of data." Similarly, the Act identifies strategic areas of policy intervention, among which (a) a certification or labeling framework for data intermediaries, and (b) measures facilitating data altruism. Yet, how to systemically manage *fairly* such emerging ecosystem of diverse actors remains an open issue.

6.2. The data republic

Based on the strengths and weaknesses of commons-inspired initiatives as detailed in Section 4.1 (Figure 1) envisions the articulation of a two-tier model aimed at guaranteeing locals' representativeness backed up by sufficient institutional and infrastructural support. Specifically, the two-tier model is composed of a Public Data Trust (PDT) with community-based Data Communes. This two-tier model

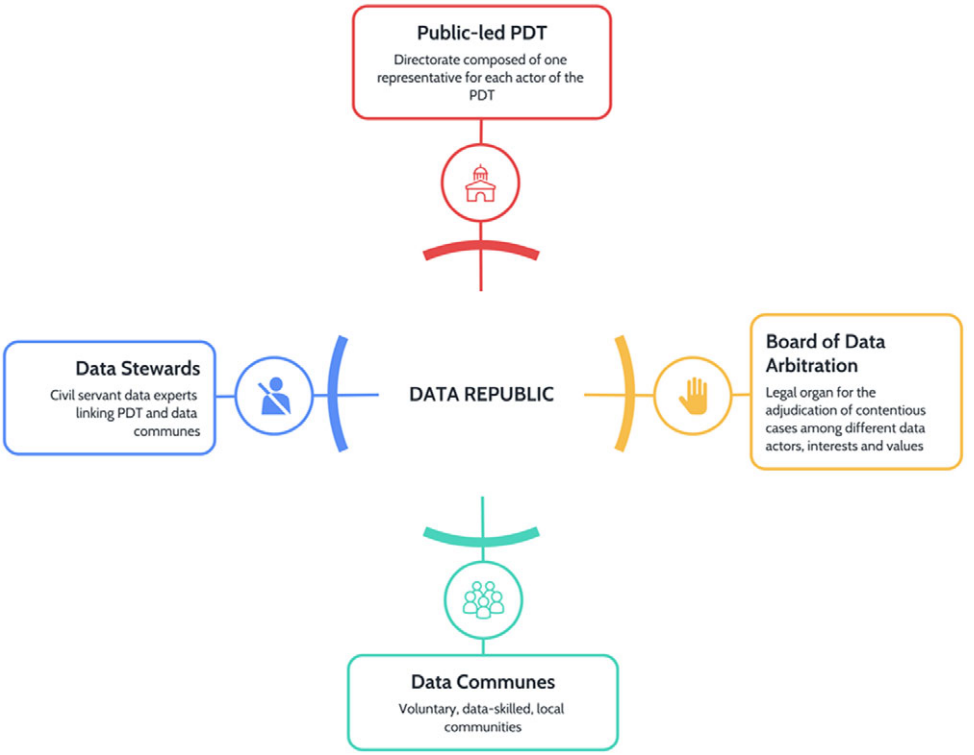

Figure 1. *Model of data republic.*

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

allows for an institutional (i.e., public-led) and infrastructural (i.e., supported by consolidated SDIs, data policies, and practices) setup to be coupled with grassroots initiatives attentive to the local dimension. A PDT (Micheli et al., 2020) "refer[s] to a model of data governance in which a public actor accesses, aggregates and uses data about its citizens, including data held by commercial entities, with which it establishes a relationship of trust." Hence, a PDT is a public-led organ which creates the conditions, under certain rules, for the commoning of data—including access, reuse, and managing—provided by a diverse array of actors: public, private, academics, citizens, and noninstitutional ones. The constitution of the PDT as an institutional organ allows avoiding the limits of most DC initiatives, which fail to scale or replicate. In fact, the PDT is an organ endowed with allocated public funding, whose actions are independent from political turnouts, and which has deliberative power concerning data management. The PDT is composed by a directorate (Lupi, 2019) composed of one representative for each data stakeholder and the directorate is renewed on a periodic basis.

As an institutional public organ, the PDT taps directly into existing infrastructures and legal frameworks and leaves room for granular data practices. Hence, the PDT works as a catalyzer for all the actors who want to contribute to the data ecosystem; as an enabler for funding and tech/legal capabilities; and a guarantor of the complying, by all actors, to the rules for the commoning of data (an OD policy should be preferred whenever possible, but this optimum can be renegotiated, via collectively agreed decisions). At the same time, however, in order to avoid the locking up of the PDT into a form of self-referentiality, which might prevent an effective participation of citizens, data communes are also envisioned.

A data commune (Susskind [2022] speaks of "mini-publics") can aggregate on a voluntary basis for having its voice heard about a specific (data-related) issue. A data commune, then, is the magnifying lens at local level of data-related issues that institutional actors do not have the flexibility to attune to. To have a data commune, the (self)identified community gathers, collects data relevant to the issue to be solved, and then asks to be formally recognized by the PDT (the data commune can also have a limited temporal existence, while the PDT ensure the durability of the data pooled). The recognition of the data commune, based on the provision of quality data, allows the data commune to become part of the PDT, with one representative. In this way, the data commune can be involved in the city's data governance—through the PDT directorate—in exchange for the contribution to the data lake of the PDT with its own indigenous data. Foreseeing that one of the barriers to data communes is the data literacies in the citizenry, their constitution is supported by pools of data stewards. This means that a data commune, even before being recognized as such, can asks the help of such figures.

Data stewards are public servant-data experts, whose role is that of mediators between the PDT and data communes. Moving beyond the corporate sector, Verhulst (2021) identifies data stewards as experts "identifying opportunities for productive cross-sector collaboration and responding pro-actively to external requests for functional access to data, insights or expertise." In this respect, data stewards are key enablers for the merging of OD frameworks and SDIs with local communities. Among data stewards' main tasks are: (a) advise the PDT, data communes, as well as other actors on data-related matters; (b) support data capability building within the public sector, as well as coordinate data literacy programs for the communities; and (c) counsel lawyers on tech-legal related matters. Data stewards should become an increasingly stabilized figure in the data republic context, and this might also mean to enforce programs for their recruitment and formation.

Lastly, representatives of the data communes, members of the PDT, and a selected number of data stewards come to constitute the board of data arbitration (renewed on a regular basis). To the extent to which the data republic enacts a fair data ecosystem where the concentration of power is to be avoided by balancing out the interests of all stakeholders involved, the board of data arbitration is meant to preserve the equilibrium of the whole data ecosystem and pursue its general interest. The board, then, works in the spirit of a jury and is responsible for adjudicating contentious issues, which can happen at various scale, based on conflicting values (e.g., individual and collective), or across various actors (e.g., between data communes or involving private and public actors). The decision of the board shall be legally binding or only consultative depending on the nature of the issue at stake. The board of arbitration shall not sanction the working of the PDT, although it can have a consultative (nonbinding) role to the PDT.

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

7. Conclusion

In this article, we introduced the concept of fair data ecosystem as an alternative to corporate-driven, state-led, *and* citizen-centric approaches to digital transformation. These approaches show limitations especially in terms of socio-economic inclusiveness, R&D diversification, and the balancing of individual and collective values. Based upon an ecosystemic understanding of fairness, a fair data ecosystem comprises roles, rules, and mechanisms to systemically take into account and, when needed, adjudicate among the data interests and values of the actors involved—beyond the quadruple helix—in order to keep the ecosystem in balance in view of common goals.

As a concretization of such fair data ecosystem, we proposed the design of a data republic. A republic is, by definition, a system that, through checks and balances, prevents the accumulation of too much unaccountable power in any actor's hands. To enable the emergence of the data republic, we advanced the coupling of a DC approach with OD frameworks and SDIs. We showed that these three regimes present complementarity features, with DC enacting participation and socio-economic inclusiveness at micro level, but lacking scalability and replicability, while OD and SDIs promote economic benefits and are consolidated at institutional and infrastructural levels, but lack the needed granularity to respond to locals' contextual needs with a collective outlook in mind.

While such coupling allows enticing institutioning and infrastructuring as the dynamic interplay of top-down and bottom-up stances, in practice this requires to disentangle on a rolling basis what the "community" and the "general interest" of a given data ecosystem are. From here, we design the data republic model as consisting of a two-tier articulation of a public-led PDT with voluntary "data communes"; a "board of data arbitration" for disentangling contentious issues on data management, and "data stewards" as public servants responsible to provide a bridge between PDT and data communes. The model emerges at the intersections of the limitations shown by OD, SDIs, and DC regimes; however, we do not claim here that the model will fix these limitations; instead, we propose to look at these regimes complementarily to move away from an understanding of data governance as targeting certain actors over others and prioritizing economic, individualistic performance over the social and collective dimensions. At stake is, above all, the ability of the model to foster *systemic* links between institutional and noninstitutional actors, as well as to negotiate between top-down and bottom-up processes. Although, in principle, a data community and general interest are fractal concepts, from a practical point of view the city is a privileged locus for testing the data republic model not only because the city is a meso-dimension linking local and (supra)national levels, but also because the city is at once a unique place of tech innovation (Jacobs, 1969) and a major target of this same innovation.

We are aware that, at present, the model of the data republic is still at a high-level of abstraction and demands not only a more fine-grained operationalization at legal and technical levels, but also a cognizant design of long-term strategies for tackling organizational issues. In fact, the model designed here needs to move beyond the page and find a practical enactment, ideally via action research, in order to test its robustness (as well as un/foreseen barriers). At this stage, however, it is already possible to indicate some policy-oriented steps to favor the enactment of such model. Notably policy efforts that maintain the city as a pivotal dimension are necessary to build (a) long-term tech-legal capacity in the public sector; (b) data literacy in citizenry; and (c) trust across institutional and noninstitutional actors. It is a whole process that needs to be fostered, and this requires mobilization of educational programs, conjoint public-private funding, and ongoing political support.

Competing interest. The authors declare no competing interests exist.

Author contribution. S.C. and B.v.L. designed the rationale of the article. S.C. wrote the article. B.v.L. provided comments and revised the article.

Data availability statement. Data availability is not applicable to this article as no new data were created or analyzed in this study.

Funding statement. This work received no specific grant from any funding agency, commercial or not-for-profit sectors.

References

Ajuntament de Barcelona (2019) Digital Transformation: City Data Commons. Available at https://ajuntament.barcelona.cat/digital/en/digital-transformation/city-data-commons (accessed 25 April 2023).

Arnstein SR (2000) A ladder of citizen participation. *Journal of the American Institute of Planners 35*(4), 216–224.

Au L and Kuuskemaa M (2019) *Social control or a fix for a non-law-abiding society? Mercator Institute for China Studies.* Available at https://www.merics.org/en/blog/social-control-or-fix-non-law-abiding-society (accessed 25 April 2023).

Balestrini M, **Rogers Y**, **Hassan C**, **Creus J**, **King M and Marshall P** (2017) A city in common: A framework to orchestrate large-scale citizen engagement around urban issues. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. New York: Association for Computing Machinery, pp. 2282–2294.

Bangratz M and Förster A (2021) *Local data and global ideas: Citymaking in times of digital transformation. PND*, 2. https://doi.org/10.18154/RWTH-2021-10411

Beckwith R, **Sherry J and Prendergast D** (2019) Data flow in the smart city: Open data versus the commons. In de Lange M and de Waal M (eds.), *The Hackable City*. Singapore: Springer, pp. 205–221.

Bendiek A and Schallbruch M (2019) *Europas dritter Weg im Cyberraum. SWP-Aktuell* 60. Available at https://www.swp-berlin.org/10.18449/2019A60/ (accessed 25 April 2023).

Budhathoki NR, **Bruce B and Nedovic-Budic Z** (2008) Reconceptualizing the role of the user of spatial data infrastructure. *GeoJournal 72*(3/4), 149–160.

Calzada I and Almirall A (2020) Data ecosystems for protecting European citizens' digital rights. *Transforming Government: People, Process and Policy 14*(2), 133–147.

Calzati S (2021) Decolonising "data colonialism" propositions for investigating the realpolitik of today's networked ecology. *Television & New Media 22*(8), 914–929.

Calzati S (2023) Federated data as a commons: A third way to subject-centric and collective-centric approaches to data epistemology and politics. *Journal of Information, Communication and Ethics in Society 21*(1), 16–29. https://doi.org/10.1108/JICES-09-2021-0097

Cardullo P and Kitchin R (2019) Smart urbanism and smart citizenship: The neoliberal logic of 'citizen-focused' smart cities in Europe. *Politics and Space C 37*(5), 813–830.

Cazacu S, **Hansen NB and Schouten B** (2020) Empowerment approaches in digital civics. In *Proceedings of OzChi '20*. https://doi.org/10.1145/3441000.3441069

Cristofari G and Helmond A (2023) The politics of platformization: Amsterdam dialogues on platform theory. *INC Theory on Demand*, 48.

Crompvoets J, **Rajabifard A**, **van Loenen B and Delgado Fernandez T** (eds.) (2008) *A Multi-View Framework to Assess Spatial Data Infrastructures*. Melbourne: Melbourne University Press.

de Lange M and de Waal M (2019) *The Hackable City: Digital Media and Collaborative City-Making in the Network Society.* Berlin: Springer Nature.

Donahoe E and Metzger M (2019) Artificial intelligence and human rights. *Journal of Democracy 30*(2), 115–126.

European Commission (2013) Powering European Public Sector Innovation: Towards a New Architecture. Available at https://ec.europa.eu/futurium/en/system/files/ged/42-public_sector_innovation_-_towards_a_new_architecture.pdf (accessed 25 April 2023).

European Commission (2014) Report on High-Value Datasets from EU Institutions. Available at https://ec.europa.eu/isa2/sites/default/files/publications/report-on-high-value-datasets-from-eu-institutions_en.pdf (accessed 25 April 2023).

European Commission (2021) Digital Europe Programme (DIGITAL). Available at https://digital-strategy.ec.europa.eu/en/news/first-calls-proposals-under-digital-europe-programme-are-launched-digital-tech-and-european-digital (accessed 25 April 2023).

European Commission (2022) European Declaration on Digital Rights and Principles for the Digital Decade. Available at https://digital-strategy.ec.europa.eu/en/library/declaration-european-digital-rights-and-principles (accessed 25 April 2023).

European Parliament and Council (2003) Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the Re-use of Public Sector Information. OJ L 345, 31/12/2003, pp. 0090–0096. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32003L0098 (accessed 25 April 2023).

European Parliament and Council (2007) Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). OJ L 108, 25.4.2007, pp. 1–14. Available at https://data.europa.eu/eli/dir/2007/2/oj (accessed 25 April 2023).

European Parliament and Council (2013a) Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 Amending Directive 2003/98/EC on the Re-use of Public Sector Information. OJ L 175, 27.6.2013, pp. 1–8. Available at https://data.europa.eu/eli/dir/2013/37/oj (accessed 25 April 2023).

European Parliament and Council (2013b) Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on Information and Communications Technology Cybersecurity Certification and Repealing Regulation (EU) No 526/2013 (Cybersecurity Act). OJ L 151e, 7.6.2019, pp. 15–69. Available at https://data.europa.eu/eli/reg/2019/881/oj (accessed 25 April 2023).

European Parliament and Council (2016) 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

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

Data, and Repealing Directive 95/46/EC (General Data Protection Regulation). OJ L 119, 4.5.2016, pp. 1–88. Available at https://data.europa.eu/eli/reg/2016/679/oj (accessed 25 April 2023).

European Parliament and Council (2018) 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. OJ L 303, 28.11.2018, pp. 59–68. Available at https://data.europa.eu/eli/reg/2018/1807/oj (accessed 25 April 2023).

European Parliament and Council (2019) 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 (Recast) (Open data directive). OJ L 172, 26.6.2019, pp. 56–83. Available at https://data.europa.eu/eli/dir/2019/1024/oj (accessed 25 April 2023).

European Parliament and Council (2020a) A European Strategy for Data. COM/2020/66 Final. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066 (accessed 25 April 2023).

European Parliament and Council (2020b) Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European Data Governance and Amending Regulation (EU) 2018/1724 (Data Governance Act). OJ L 152, 3.6.2022, pp. 1–44. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868 (accessed 25 April 2023).

European Parliament and Council (2022) Proposal for a Regulation of the European Parliament and of the Council on Harmonised Rules on Fair Access to and Use of Data (Data Act) COM/2022/68 Final. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A68%3AFIN (accessed 25 April 2023).

Feinberg A, **Ghorbani A and Herder P** (2021) Diversity and challenges of the urban commons: A comprehensive review. *International Journal of the Commons 15*(1), 1–20.

Fu X, **Woo WT and Hou J** (2016) Technological innovation policy in China: The lessons, and the necessary changes ahead. *Economic Change and Restructuring 49*, 139–157.

Genin A, **Tan J and Song J** (2021) State governance and technological innovation in emerging economies: State-owned enterprise restructuration and institutional logic dissonance in China's high-speed train sector. *Journal of International Business Studies 52*, 621–645.

Glasze G, **Cattaruzza A**, **Douzet F**, **Dammann F**, **Bertran MG**, **Bômont C**, **Braun M**, **Danet D**, **Desforges A**, **Géry A**, **Grumbach S**, **Hummel P**, **Limonier K**, **Münßinger M**, **Nicolai F**, **Pétiniaud L**, **Winkler J and Zanin C** (2023) Contested spatialities of digital sovereignty. *Geopolitics 28*, 919–958.

Han S, **Liu H and Lin Y** (2019) Measurement of the innovation efficiency of the hi-tech industry in China and its influencing factors. *International Journal of Sustainable Development and Planning 15*(3), 277–286.

Hasselbalch, G. (2021). A framework for a data interest analysis of artificial intelligence. First Monday. https://firstmonday.org/ojs/index.php/fm/article/download/11091/10168

Hausstein A and Zheng C (eds.) (2018) *Industrie 4.0/Made in China 2025. Gesellschaftswis-senschaftliche Perspektiven auf Digitalisierung in Deutschland und China*. Karlsruhe: KIT.

Healey P (1997) *Collaborative planning*. London: McMillan.

High-Level Expert Group on Artificial Intelligence (2019) Ethics Guidelines for Trustworthy AI. Available at https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (accessed 25 April 2023).

Hummels P, **Braun M and Dabrock P** (2021) Own data? Ethical reflections on data ownership. *Philosophy & Technology 34*, 545–572.

Huybrechts L, **Benesch H and Geib J** (2017) Institutioning: Participatory design, co-design and the public realm. *CoDesign 13*(3), 148–159.

Innes JE and Booher DE (2015) A turning point for planning theory? Overcoming dividing discourses. *Planning Theory 14*(2), 195–213.

Jacobs, J. (1969). *The economy of cities*. New York: Random House.

Jarke M, **Otto B and Ram S** (2019) Data sovereignty and data space ecosystems. *Business & Information Systems Engineering 61* (5), 549–550.

Jasperneite J (2012) Was hinter Begriffen wie Industrie 4.0 steckt. *Computer & Automation*, 19.

Jiang M and Fu K-W (2018) Chinese social media and big data: Big data, big brother, big profit? *Policy and Internet 10*(4), 372–392.

Kalpokas I (2022) Posthuman urbanism: Datafication, algorithmic governance and Covid-19. In Bobic N and Haghighi F (eds.), *The Routledge Handbook of Architecture, Urban Space and Politics, Volume I*. London: Routledge, pp. 496–508).

Kostakis V, **Niaros V**, **Dafermos G and Bauwens M** (2015) Design global, manufacture local: Exploring the contours of an emerging productive model. *Futures 73*, 126–135.

Kozminski JJ, **Zoboli L and Nemitz P** (2021) Embedding European values in data governance: A case for public data commons. *Internet Policy Review 10*(3), 1–29.

Kummitha RKR (2020) Why distance matters: The relatedness between technology development and its appropriation in smart cities. *Technological Forecasting and Social Change 157*, 120087.

Latour B (2004) *The Politics of Nature: How to Bring the Sciences into Democracy* (C. Porter, trans.). Cambridge, MA: Harvard University Press.

Lee K F (2019) *AI Superpowers: China, Silicon Valley, and the New World Order*. Boston: Houghton Mifflin.

Lee MSA, **Floridi L and Singh J** (2021) Formalising trade-offs beyond algorithmic fairness: Lessons from ethical philosophy and welfare economics. *AI & Ethics 1*, 529–544.

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

Lee MSA and Singh J (2021) The landscape and gaps in open source fairness toolkits. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. New York: Association for Computing Machinery, pp. 1–13.

Lippert B and Perthes V (eds.) (2020) Strategic rivalry between United States and China. Causes, trajectories, and implications for Europe. *SWP Research Paper* 4. https://doi.org/10.18449/2020RP04

Ludwig T, **Pipek V and Tolmie P** (2018) Designing for collaborative infrastructuring: Supporting resonance activities. In *Proceedings of the ACM on Human-Computer Interaction*. New York: Association for Computing Machinery. https://doi.org/10.1145/3274382

Lupi L (2019) City data plan: The conceptualisation of a policy instrument for data governance in smart cities. *Urban Science 3*(3), 91. https://doi.org/10.3390/urbansci3030091

Lupi L, **Antonini A**, **De Liddo A and Motta E** (2020) Actionable open data: Connecting city data to local actions. *Journal of Community Informatics 16*, 3–25.

Martin C, **Evans J and Karvonen A** (2018) Smart and sustainable? Five tensions in the vision and practices of the smart-sustainable city in Europe and North America. *Technological Forecasting of Social Change 133*, 269–278.

Masser I (1999) All shapes and sizes: The first generation of national spatial data infrastructures. *International Journal of Geographical Information Science 13*(1), 67–84.

Mergel I, **Kleibrink A and Srvik J** (2018) Open data outcomes: U.S. cities between product and process innovation. *Government Information Quarterly 35*(4), 622–632.

Micheli M, **Ponti M**, **Craglia M and 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

Monge F, **B-H CL Initiative**, **Barns S**, **Kattel R and Bria F** (2022) A New Data Deal: The Case of Barcelona. Available at https://www.ucl.ac.uk/bartlett/public-purpose/sites/bartlett_public_purpose/files/new_data_deal_barcelona_fernando_barns_kattel_and_bria_18_feb.pdf (accessed 25 September 2023).

Morozov E and Bria F (2018) *Rethinking the Smart City. Democratizing Urban Technology*. New York: Rosa Luxemburg Foundation.

Mulder AE, **Wiersma G and van Loenen B** (2020) Status of national open spatial data infrastructures: A comparison across continents. *International Journal of Spatial Data Infrastructures Research 15*, 56–87.

Mulder I and Kun P (2019) Hacking, making, and prototyping for social change. In de Lange M and de Waal M (eds.), *The Hackable City*. Singapore: Springer, pp. 225–238.

Open Knowledge Foundation (2013) Open Data Definition. Available at https://okfn.de/en/themen/open_data/ (accessed 25 April 2023).

Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

Pew Research Center (2020) Many Experts Say Digital Disruption Will Hurt Democracy. Available at https://www.pewresearch.org/internet/2020/02/21/many-tech-ex-perts-say-digital-disruption-will-hurt-democracy/ (accessed 25 April 2023).

Pohle J and Thiel T (2020) Digital sovereignty. *Internet Policy Review 9*(4), 1–19.

Raymond A and Kouper I (2023) Open governments, open data: Moving toward a digital commons framework. In Frischmann B, Madison M and Sanfilippo M (eds.), *Governing Smart Cities as Knowledge Commons*. Cambridge: Cambridge University Press, pp. 58–80.

Roberts H, **Cowls J**, **Morley J**, **Taddeo M**, **Wang V and Floridi L** (2021) The Chinese approach to artificial intelligence: An analysis of policy, ethics, and regulation. *AI & Society 36*, 59–77.

Rochel J (2021) Ethics in the GDPR: A blueprint for applied legal theory. *International Data Privacy Law 11*(2), 209–223.

Sanfilippo M and Frischmann B (2023) A proposal for principled decision-making: Beyond design principles. In Frischmann B, Madison M and Sanfilippo M (eds.), *Governing Smart Cities as Knowledge Commons*. Cambridge: Cambridge University Press, pp. 295–308.

Schneider I (2020) Democratic governance of digital platforms and artificial intelligence?: Exploring governance models of China, the US, the EU and Mexico. *JeDEM–eJournal of eDemocracy and Open Government 12*(1), 1–24.

Scott G and Rajabifard A (2017) Sustainable development and geospatial information: A strategic framework for integrating a global policy agenda into national geospatial capabilities. *Geo-spatial Information Science 20*(2), 59–76.

Shkabatur J (2019) The global commons of data. *Stanford Technology Law Review 22*, 354–411.

Smuha NA (2021) Beyond the individual: Governing AI's societal harm. *Internet Policy Review 10*(3), 1574. https://doi.org/10.14763/2021.3.1574

SPIDER Consortium (2021) Exploring and Defining the Open SDI Concept. Available at https://sdispider.eu/wp/wp-content/uploads/2020/11/IO2_29_10_2020.pdf (accessed 25 September 2023).

Star S and Ruhleder K (1996) Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research 7*(1), 111–134.

Streeck W and Thelen K (2005) *Beyond Continuity: Institutional Change in Advanced Political Economies*. New York: Oxford University Press.

Sun P (2007) Is the state-led industrial restructuring effective in transition China? Evidence from the steel sector. *Cambridge Journal of Economics 31*(4), 601–624.

Susskind J (2022) *The Digital Republic: On Freedom and Democracy in the 21st Century*. London: Bloomsbury Publishing.

https://doi.org/10.1017/dap.2023.18 Published online by Cambridge University Press

Tannier C and Thomas I (2013) Defining and characterizing urban boundaries: A fractal analysis of theoretical cities and Belgian cities. *Computers, Environment and Urban Systems 41*, 234–248.

Taylor L (2021) Public actors without public values: Legitimacy, domination and the regulation of the technology sector. *Philosophy & Technology 34*, 897–922.

Toots M (2019) Why e-participation systems fail: The case of Estonia's Osale.ee. *Government Information Quarterly 36*, 546–559.

Turner F (2017) Don't be evil. On utopias, frontiers, and brogrammers. *Logic* 3, December 01. https://logicmag.io/03-dont-be-evil/ (accessed 25 April 2023).

Valli Buttow C and Weerts S (2022) Public sector information in the European Union policy: The misbalance between economy and individuals. *Big Data & Society 9*(2). https://doi.org/10.1177/20539517221124587

van Dijck J, **Nieborg D and Poell T** (2019) Reframing platform power. *Internet Policy Review 8*(2), 1–18.

van Loenen B (2006) *Developing Geographic Information Infrastructures: The Role of Information Policies*. Amsterdam: IOS Press.

van Loenen B, **Zuiderwijk A**, **Vancauwenberghe G**, **Lopez-Pellicer FJ**, **Mulder I**, **Alexopoulos C**, **Magnussen R**, **Saddiqa M**, **Dulong de Rosnay M**, **Crompvoets J**, **Polini A**, **Re B and Casiano Flores C** (2021) Towards value-creating and sustainable open data ecosystems: A comparative case study and a research agenda. *JeDEM–eJournal of eDemocracy and Open Government 13*(2), 1–27.

Vancauwenberghe G, **Valečkaitė K**, **van Loenen B and Welle Donker F** (2018) Assessing the openness of spatial data infrastructures (SDI): Towards a map of open SDI. *International Journal of Spatial Data Infrastructures Research 13*, 88–100.

Vancauwenberghe G and van Loenen B (2018) Exploring the emergence of open spatial data infrastructures: Analysis of recent developments and trends in Europe. In Saeed S, Ramayah T and Mahmood Z (eds.), *User Centric E-Government. Challenges and Opportunities*. New York: Springer International Publishing, pp. 23–45.

Verhulst S (2021) Reimagining data responsibility: 10 new approaches toward a culture of trust in re-using data to address critical public needs. *Data & Policy 3*, e6. https://doi.org/10.1017/dap.2021.4

Verhulst S, **Young A**, **Zahuranec AJ**, **Aaronson SA**, **Calderon A and Gee M** (2020) *Third wave of open data. GovLab.* https://opendatapolicylab.org/images/odpl/third-wave-of-opendata.pdf (accessed 25 April 2023).

Viitanen J and Kingston R (2014) Smart cities and green growth: Outsourcing democratic and environmental resilience to the global technology sector. *Environment and Planning A 46*(4), 803–819.

Viljoen S (2021) A relational theory of data governance. *Yale LJ 131*, 573.

von der Leyen U (2019) A Union that Strives for More: My Agenda for Europe. Available at https://ec.europa.eu/info/sites/default/files/political-guidelines-next-commission_en_0.pdf (accessed 25 April 2023).

Voss G and Pernot-Leplay E (2023) China data flows and power in the era of Chinese big tech. *Northwestern Journal of International Law & Business 44*, 4393008. https://doi.org/10.2139/ssrn.4393008

Welle Donker F and van Loenen B (2017) How to assess the success of the open data ecosystem? *International Journal of Digital Earth 10*(3), 284–306.

Welle Donker F, **van Loenen B and Bregt AK** (2016) Open data and beyond. *ISPRS International Journal of Geo-Information 5* (4), 1–16.

Wen Y (2021) *The Huawei Model: The Rise of China's Technology Giant*. Champaign: University of Illinois Press.

Wolff A, **Gooch D**, **Cavero J**, **Rashid U and Kortuem G** (2019) Removing barriers for citizen participation to urban innovation. In de Lange M and de Waal M (eds.), *The Hackable City.* Singapore: Springer, pp. 153–168.

Yu Y, **Yu J**, **Pan X and Stough R** (2017) Governance and the China innovation economy. *Asia-Pacific Journal of Regional Science 1*(1), 63–84.

Zeng DZ (2017) Measuring the effectiveness of the Chinese innovation system: A global value chain approach. *International Journal of Innovation Studies 1*(1), 57–71.

Cite this article: Calzati S and van Loenen B (2023). A fourth way to the digital transformation: The data republic as a fair data ecosystem. *Data & Policy*, 5: e21. doi:10.1017/dap.2023.18