Delft University of Technology

Back to Roots: Reconsidering Data-As-Resource to Introduce a Commons-Based Modulating Approach to Open Data

Calzati, S.

DOI
[10.5281/zenodo.8070068](https://doi.org/10.5281/zenodo.8070068)

Publication date
2022

Document Version
Final published version

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

International Conference
on Open Data
(ICOD 2022)

BOOK OF ABSTRACTS

International Conference on Open Data: Open Data Challenges and Opportunities in Times of Crisis and Growth (ICOD 2022)
November 28th – December 2nd 2022, Zagreb, Croatia

ICOD 2022
International Conference on Open Data:
Open Data Challenges and Opportunities in Times of Crisis and Growth
November 28th – December 2nd, 2022, Zagreb, Croatia

Organized by:

Faculty of Law, University of Zagreb

and

Faculty of Agriculture, University of Zagreb
Faculty of Geodesy, University of Zagreb
Faculty of Electrical Engineering and Computing, University of Zagreb
Faculty of Organisation and Informatics, University of Zagreb
Delft University of Technology, Netherlands
University of Aegean, Greece

within the consortium of the project:
Twinning open data operational (TODO) – 857592

BACK TO ROOTS: RECONSIDERING DATA-AS-RESOURCE TO INTRODUCE A COMMONS-BASED MODULATING APPROACH TO OPEN DATA

Stefano Calzati[1*]

[1] Faculty of Architecture & the Built Environment, Delft University of Technology, Netherlands

*correspondence E-mail: s.calzati@tudelft.nl

Keywords: data commons; data ecosystem; data resource; open data; openness modulation, EU

This paper takes the lead from the acknowledged ongoing mismatch between open data provision and actual open data demand by users, both public and private. To tackle such mismatch, the paper suggests remodulating data openness through a commons approach, which seeks to account for all data actors' needs, while maintaining their interests in balance. In other words, data commons enacts an ecosystemic approach by default to data provision and demand, whereby data actors and their relations are mutually co-dependent. Yet, the application of the concept of the commons to data is not straightforward. This is due to data's hybrid nature, which requires reconsidering the normative understanding of data as a resource, if we really want to develop a governance model that is beneficial to all actors *and* the whole data ecosystem. Building on recent literature in this direction, the paper will briefly outline the pillars of a commons-based data ecosystem.

Since 2003 the EU (European Parliament, 2003; 2013; 2016; 2018; 2019; European Commission 2013; 2014) has released directives about the re-use of public sector information with the goal to have this information published as open data. Recently, the Data Governance Act (European Commission, 2020) and the proposed Data Act (European Commission, 2022) constitute two further pillars in this direction aimed at boosting business-to-government (B2G), citizen-to-business (C2B) and business-to-business (B2B) data sharing. While these directives signal an increasing drive towards open data and the fostering of a data-inclusive ecosystem, limitations remain. In fact, today open data are still under-exploited (Lupi et al., 2020) due also to a misalignment between the provision of data and the actual information needs of local actors. It is necessary to put equal emphasis on data supply and on data demand, for the timely release of appropriate data (Welle Donker et al., 2016). Moreover, open data initiatives have so far targeted national and supranational scales, while much data reside at local level (Verhulst, 2021). Hence, while open data represent a key institutionalized enabler to digital innovation, what is still missing are mechanisms and practices connecting (supra)nation and local levels, thus favoring the matching between provision and demand of data across scales. This is why scholars have encouraged authorities at various levels to engage citizens and promote the formation of stakeholder communities around open data (Mergel et al., 2018).

To tackle this mismatch, it is advanced here the possibility to remodulate the idea of data openness through a commons approach that can contribute to realize a sustainable open data ecosystem (van

Loenen et al., 2021). When speaking of (open) data – and more generally of data ecosystem – one should always ask: Open *how*? For *whom*? According to which *rules* and *values*? At stake is an issue not much of ownership, but control of data, latter of which is a concept open – much more than the former – to modulation. Openness, in other words, is a thick concept that lends itself and demands to be articulated on a contextual and rolling basis. As Hummels et al. (2021) note in the context of data use and reuse: "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." Concretely, this means that while the EU's striving for data openness remains the polar star, this represents an optimum to which to tend (rather than a default standard) and which requires continuous modulation to attune to and harness from all data actors' needs and provisions.

As a regime for the managing of resources, the concept of the commons can prove fruitful in this scenario. Originally, the "commons" referred to common-pool resources (CPRs) – such as fisheries or forests – characterized by non-excludability and rivalry, meaning respectively that: 1) it is difficult to forbid access and use of CPRs to any potential beneficiary; and 2) the use of CPRs depletes them and reduces further use by others. Ostrom (1990) shows 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). This characterization contains the idea that the commons exists to the extent there is a *commoning practice* (de Angelis, 2007) that conceives and manages the resource as a commons. The trading mark of such practice is to be non-appropriative by default (knowledge, technology, outputs are not owned, in the commercial sense of the term, but controlled and summoned up); collaborative by design (it considers all actors as integral to the ecosystem's flourishing), and collectively sustainable in its goals (indeed, commons for the community).

When it comes to data, data commons characterizes a regime in which actors join forces in the collection, pooling, and use of data (and infrastructures) subservient to the delivery of services for the whole community (de Lange and de Waal, 2019). As Hess and Ostrom (2007) explain, technology plays a crucial role in applying the commons to a given resource: "[t]his ability [of technology] to capture the previously uncapturable creates a fundamental change in the nature of the resource, with the resource being converted from a nonrivalrous, nonexclusive public good into a common-pool resource that needs to be managed, monitored, and protected, to ensure sustainability and preservation." This means that as soon as a (new) technology seizes or creates a resource, this can effectively be managed as a commons.

In this respect, data commons initiatives (Morozov and Bria, 2018) aim to counteract and/or repurpose the centralized ownership and use of data – either by tech companies or governments – by giving these back to people, with the goal to foster sustainable collective data practices. Data commons initiatives truly reinserts people into the data ecosystem and allow them to co-develop tech solutions with a collective socio-economic (and environmental) outlook in sight. This is why to redesign open data initiatives through a commons-driven approach might prove as a viable path to tackle the mismatch between open data provision and demand.

To do so, however, it is necessary to go back to roots and rethink our understanding of data, insofar as, while the coupling of (open) data with the commons is potentially fruitful, it is not straightforward. This is due to the current understanding of data as a resource, which is at the basis of normative approaches to data governance, including the EU's. Despite the consolidated metaphor of data as the "new oil", data are different from traditional natural resources. On this point, Prainsack (2019) notes that "although digital data clearly have material components, their materiality is of a very different kind than the physical resources." From here, it does not surprise to find both advocates of data's global nature (Shkabatur, 2019) as well as defendants of data's local rootedness (Loukissas, 2019). These positions are indeed symptomatic of the difficulty to consider data as a resource in the traditional sense of the term, that is, as a material to be found "out there" and capitalized upon. Once we dig into the nature of data-as-resource, a series of consequences derive about how to properly reconceptualize and manage data.

Differently from natural resources, data do not pre-exist in nature. Instead, data are a fully artificial (human and/or tech-created) *construct* that exists at the very moment a certain (sociotechnical) process is enacted to collect *certain* information (this is also why between information and data there is never a straightforward link, but rather a fundamental qualitative gap). This leads to suggest that data-as-resource is unique in that it manifests an *entangled nature*: if one stresses the *informational* constituency of data, then data are a virtual entity and are potentially distributable globally; if one stresses the *technical* constituency of data (from collection to storing and use), then data are material entities whose circulation can be hindered in many ways (Bates, 2018). This is why it makes little sense to ask whether data (as a resource) is scarce or abundant: data is neither scarce nor abundant – it is its use-in-context that (should) always dictate its provision. As an example, take the General Data Protection Regulation (European Parliament, 2016): a strict interpretation of the GDPR (as the "law of everything", see Purtova, 2018) signals an unbalance towards the informational side of data; a too loose interpretation of the GDPR foregrounds the technical side of data. To be sure, the entangled nature of data is also at the basis of the conflicting approaches to data ownership (Hummels et al., 2021), depending on which side is given prominence.

Therefore, back to the commons, it is somewhat misleading to ask whether data-as-resource is non/excludable and/or non/rivalrous. On the one hand, being informational, data can be accessed (in principle) by anyone, thus making them non-excludable; being technical, their access can be prevented, thus making them an excludable resource; on the other hand, being virtual, data are non-rivalrous, while as a technical artefact that contains certain information they are rivalrous. The managing of data as a commons, then, requires a paradigmatic shift in the way we think about data, passing from data as a "thing" to data as a (sociotechnical) *process*: data are always created under certain conditions, used for certain purposes, in certain contexts, and with certain results. Governance mechanisms must be designed to negotiate between the two constituencies of data and/or disentangle and give priority to either one of the two. To really enact a match between provision and demand of open data, data governance needs to move away from approaches that either target certain actors over others – e.g., citizens, public actors, private actors – or prioritize one value over others – oftentimes economic competitiveness over social inclusion or sustainability – to rather adopt an ecosystemic standpoint (Jarke et al., 2019) which, by definition, is irreducible to any of its components or relations and requires an integrated approach to find a balance of/for

the whole. In this respect, recent literature on data commons can provide a valuable starting point for advancing an ecosystemic remodulation of open data.

In her work on CPRs, Ostrom (1990) identified eight principles that allow for an effective self-management of the commons: 1) well-established boundaries of the resource; 2) rules that are context-sensitive; 3) mechanisms of collective participation for the modification of the rules; 4) these mechanisms need to be respected by external actors; 5) monitoring system for the compliance to the rules; 6) mechanisms for arbitration and sanctioning; 7) recognized and easy mechanisms for appeal; and 8) tiered management for large resources. Bloom et al. (2021) have delved into Ostrom's principles suggesting how they might be transposed in the context of data initiatives. In so doing, these authors outline guidelines concerning the governance that such commons-based data initiatives might take. If we apply to this work an ecosystemic vision, then the remodulation of open data as a commons gets inscribed into a framework whose key feature is to maintain an equilibrium among all the principles. Notably, it is possible to identify three core pillars as enablers of such equilibrium: a public-led data trust (PDT) coupled with voluntary citizen-pooled "data communes"; processes of data stewardship, and processes of data arbitration. This three-folded articulation follows up on data republicanism (Susskind, 2022), as an approach to data governance which prevents the concentration of unaccountable data power, favoring its distribution according to mutual checks and balances.

Principle 1 is the most problematic because it taps directly into the nature of the resource (i.e. data) alongside its supporting infrastructure. While data always belong to a given community/context (Bechwith et al., 2019), such community/context is entangled with both the infrastructure that supports the data and the legal framework that defines their global flow. More so: the three – data, infrastructure/legal framework, community – are always redefining poles of a whole ecosystem in the making. And it is this ecosystem in the making that needs to be "commoned". This is why the release alone of data as open data is not enough for fostering a sustainable data ecosystem; rather, rules and mechanisms are needed in the direction of a fit-for-context modulation of the control of data-as-resource (principle 2).

Principles 3 and 4 can be enacted by the constitution of a PDT which capitalizes on the institutionalization of open data, while also guaranteeing accountability and participation via a systematic opening towards diverse stakeholders. In the words of Micheli et al. (2020) a PDT "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 creates the conditions, under certain rules, for the commoning of data provided by a diverse array of actors. Yet, to avoid the locking up of the PDT in a form of institutional self-referentiality, data communes are also envisioned. A data commune is composed of citizens who aggregate on a voluntary basis for having their voice heard about a specific (data-related) issue. To have a data commune, the (self)identified group gathers, collects data relevant to the issue to be solved, and then asks to be formally recognized by the PDT. The recognition of the data commune, based on the provision of quality data, allows the data commune to become part of the PDT.

Since barriers to data commons initiatives often come in the forms of limited data literacy and tech capabilities (Monge et al., 2022), the constitution of data communes is supported through

institutionalized processes of data stewardship (principle 8). 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 linking the PDT with data communes, as well as for contributing to data literacies in citizenry and tech-legal capabilities in the public sector.

Lastly, in order to concretize principles 5, 6 and 7, it is possible to envision processes of data arbitration, in the spirit of juries responsible for deciding upon contentious issues about data use, which can happen at various scales and across various actors. Concerning both data stewardship and data arbitration, it is worth noting that since the Open Data directive (European Parliament, 2013), the EU has acknowledged the need to enforce such processes. This was reaffirmed in latest documents: about processes of data arbitration, the Data Governance Act (European Commission, 2020) states that "any natural or legal person affected by a decision of a public sector body shall have the right to an effective judicial remedy against such decision", while the proposed Data Act (European Commission, 2022) speaks of "settlement bodies" to ensure "alternative ways of resolving domestic and cross-border disputes that arise in connection with making data available." About data stewardship, the Data Governance Act 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. (…) The competent body or bodies shall have adequate legal and technical capacities and expertise to be able to comply with relevant Union or national law." Moreover, member states will have to set up "structures to support public sector bodies with technical means and legal assistance." Future research will need to explore how to properly design these processes of stewardship and arbitration to make sure that they cut across scales – from the local to the European level – and contexts, i.e. facilitating cross-boundary collaboration among member states.

Altogether, PDT+data communes, processes of data stewardship and data arbitration constitute an ideal modelling (**Figure 1**) of Ostrom's 8 principles for the managing of CPRs, applied to data initiatives. While such modelling allows, in theory, for a contingent and context-based modulation of open data, thus favoring the matching between data provision and demand, the model requires testing in practice on both open data and data commons initiatives to identify barriers and enablers to its enactment.
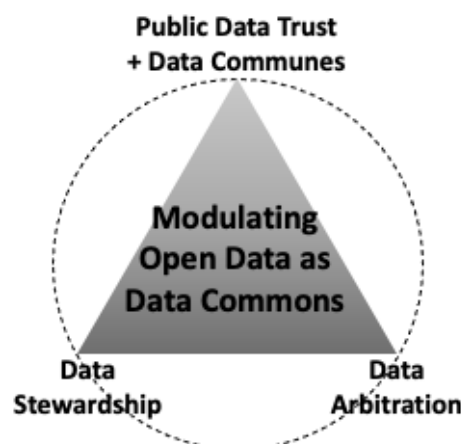
Figure 1. Modelling of open data as data commons

References

Beckwith, R., Sherry, J., & Prendergast, D. (2019). Data flow in the smart city: Open data versus the commons. In de Lange, M., de Waal, M. (Eds.) The hackable city: Digital media and collaborative city-making in the network society 205-221. Singapore: Springer.

Bloom, G., Raymond, A., Tavernier, W., Siddarth, D., Motz, G., Dulong de Rosnay, M., Ruhaak, A. (2021). A practical framework for applying Ostrom's principles to data commons governance. Available at https://foundation.mozilla.org/en/blog/a-practical-framework-for-applying-ostroms-principles-to-data-commons-governance/

de Angelis, M. (2017). Omnia sunt communia: On the commons and the transformation to postcapitalism. London: Bloomsbury Publishing.

de Lange, M., de Waal, M. (2019). The hackable city: Digital media and collaborative city-making in the network society. Singapore: Springer.

European Commission. (2022). Data Act: Proposal for a Regulation on harmonised rules on fair access to and use of data. Available at https://digital-strategy.ec.europa.eu/en/library/data-act-proposal-regulation-harmonised-rules-fair-access-and-use-data

European Commission. (2020). Data governance act. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0767

European Commission. (2014). Report on high-value datasets from EU institutions. Available at http://ec.europa.eu/isa/documents/publications/report-on-high-value-datasets-from-eu-institutions_en.pdf

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

European Parliament (2019). Open data and the re-use of public sector information. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L1024&from=EN

European Parliament (2018). Framework for the free flow of non-personal data in the European Union. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807&from=EN

European Parliament (2016). General data protection regulation. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02016R0679-20160504&from=EN

European Parliament (2003). Directive on the re-use of public sector information. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32003L0098&from=en

Feinberg, A., Ghorbani, A., Herder, P. (2021). Diversity and challenges of the urban commons: A comprehensive review. International Journal of the Commons 15(1): 1-20. http://doi.org/10.5334/ijc.1033

Hess, C., Ostrom, E. (2007). Understanding Knowledge as a Commons. Baston: MIT Press.

Hummels, P., Braun, M., Dabrock, P. (2021). Own data? Ethical reflections on data ownership. Philosophy & Technology 34: 545–572. https://doi.org/10.1007/s13347-020-00404-9

Jarke, M., Otto, B., Ram, S. (2019). Data sovereignty and data space ecosystems. Business & Information Systems Engineering 61(5): 549-550. https://doi.org/10.1007/s12599-019-00614-2

Loukissas, Y. A. (2019). All data are local: Thinking critically in a data-driven society. Boston: MIT Press.

Lupi, L., Antonini, A., De Liddo, A., Motta, E. (2020). Actionable open data: Connecting city data to local actions. Journal of Community Informatics 16: 3-25. https://doi.org/10.15353/joci.v16i0.3492

Mergel, I., Kleibrink, A., Srvik, J. (2018). Open data outcomes: U.S. cities between product and process innovation. Government Information Quarterly 35(4): 622–632. https://doi.org/10.1016/j.giq.2018.09.004

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

Monge, F., Barns, S., Kattel, R., Bria, F. A new data deal: The case of Barcelona. UCL Institute for Innovation and Public Purpose. 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

Morozov, E., Bria, F. (2018). Rethinking the smart city. Democratizing Urban Technology. New York: Rosa Luxemburg Foundation.

Ostrom, E. (1990). Governing the commons: The evolution of institutions for collective action. Cambridge: Cambridge University Press.

Purtova, N. (2018). The law of everything. Broad concept of personal data and future of EU data protection law. Law, Innovation and Technology 10(1): 40-81. https://doi.org/10.1080/17579961.2018.1452176

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

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

Susskind, J. (2022). Digital technology demands a new political philosophy. Available at https://www.noemamag.com/digital-technology-demands-a-new-political-philosophy/

van Loenen, B., Zuiderwijk, A. ., Vancauwenberghe, G., Lopez-Pellicer, F. J., Mulder, I., Alexopoulos, C., Magnussen, R., Saddiqa, M., Dulong de Rosnay, M., Crompvoets, J., Polini, A., Re, B., 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. https://doi.org/10.29379/jedem.v13i2.644

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

Welle Donker, F. van Loenen, B., Bregt, A. K. (2016). Open data and beyond. ISPRS International Journal of Geo-Information 5(4): 48. https://doi.org/10.3390/ijgi5040048