European Journal of Risk Regulation (2025), **16**, 114–132
doi:10.1017/err.2024.73

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Data-Driven Policy Making and Its Impacts on Regulation: A Study of the OECD Vision in the Light of Data Critical Studies

Clarissa Valli Buttow

Swiss School of Public Administration, University of Lausanne, Lausanne, Switzerland
Email: clarissa.vallibuttow@unil.ch

Abstract

While the private sector is leading the way in digital transformation, the public sector is perceived as missing the opportunity. The way forward in this transformation is pointed out as adopting data-driven technologies as drivers of public policies. This will undoubtedly transform the regulatory process. To better understand how these technologies would intervene in the policy cycle and the effect of this transformation on regulation, this paper studies the work of the Organisation for Economic Co-operation and Development on the matter. Adopting the public policy cycle as an analytical lens and illuminated by the insights of Critical Data Studies, this paper deepens the understanding of how these technologies could change the policy cycle and bring together steps previously conceived as subsequential stages. This transformation can trigger a real disruption of the regulatory process as we currently know it.

Keywords: better regulation; data-driven policymaking; regulatory process

I. Introduction

We currently see a race for artificial intelligence (AI) and data-driven innovation that crosscuts the public and private sectors. Scholarship, policymakers and the general public expect from technological developments the solution for social problems and economic sustainable development.[1] This trend is also present in the public sector, which appears to be hanging behind concerning innovative and technological solutions.[2] The key to unlocking innovation is pointed out as being an increased use of data and data-driven innovation in the public sector throughout government activity, in what has been called a whole-of-government approach.[3] The fact that technology and innovation are "data-driven" refers to the fact that both embrace, at their core, processes of data collection (or extraction), processing and/or analysis.[4] The expression serves as an umbrella term that

[1] J Berryhill et al. "Hello, World : Artificial intelligence and its use in the public sector," *Documents de travail de l'OCDE sur la gouvernance publique*, n° 36, Éditions OCDE, (2019 Paris), p 134.

[2] Kevin C. Desouza, Gregory S. Dawson and Daniel Chenok, "*Designing, developing, and deploying artificial intelligence systems: Lessons from and for the public sector*" (2020) 63 Business Horizons 205 doi: 10.1016/j.bushor.2019.11.004.

[3] Benjamin Welby and Ethel Hui Yan Tan, "*Designing and delivering public services in the digital age*" (2022) OECD Going Digital Toolkit Notes, n° 22, Paris. doi: 10.1787/e056ef99-en.

[4] Kayan Fukuda, "Science, technology and innovation ecosystem transformation toward society 5.0" (2020) 220 International Journal of Production Economics; Saida Sultana, Shahriar Akter, Elias Kyriazis, Samuel Fosso

conveys the idea of data at the core independently of the technology itself (it can, for example, refer to artificial intelligence powered by data, such as large language models, and to different techniques, such as machine learning algorithms or deep learning). First viewed as simple tools in public service delivery, the scope of deployment of technologies, specifically artificial intelligence powered by data-driven technologies, is growing fast. Some authors have pointed out that data-driven technologies would be able to improve public policy, enabling more efficient, effective and responsive government.[5] Nonetheless, other authors adopt a more cautious position, warning about the risks of the wide adoption of a data-driven approach to public policies.[6] All in all, in recent years, there has been an increasing appeal to develop and deploy data-driven technologies throughout the policy cycle, from the design and formulation of public policies to their implementation and evaluation.[7]

The transformation of the public policy cycle by data-driven technology will also engender a transformation of the legislative and regulatory process. Integrating data-driven technologies into the policy cycle is highly relevant from the perspective of legislation and regulation, given that the policy cycle is intertwined with the law and rulemaking processes.[8] Indeed, the regulatory and legal frameworks structure the policy cycle, while new rules or a new regulatory framework can result from a public cycle.[9] There is, therefore, a reciprocal link between policy and regulation, and the transformation of one process will impact the other.[10]

The literature in public policy and public administration has been mapping and investigating the effects of data-driven technologies throughout the policy cycle.[11] Likewise, interdisciplinary literature has investigated the diverse effects of the so-called algorithmic regulation in different domains of public action.[12] Yet, the study on the impacts of the broad adoption of data-driven technology throughout the policy cycle on

Wamba, "Architecting and Developing Big Data-Driven Innovation (DDI) in the Digital Economy" (2021) 29 Journal of Global Information Management 165–187.

[5] Helen Margetts and Cosmina Dorobantu, "Rethink government with AI: Policymakers should harness data to deliver public services that are responsive, efficient and fair, urge'(2019) 568 Nature 163 doi: 10.1038/d41586-019-01099-5; David Valle-Cruz and others, "Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation" (2020) 37 Government Information Quarterly 101509 doi: 10.1016/j.giq.2020.101509; Alberto Alemanno, "Big Data for Good: Unlocking Privately-Held Data to the Benefit of the Many" (2018) 9 European Journal of Risk Regulation 183, https://doi.org/10.1017/err.2018.34.

[6] Sarah Giest, "Big data for policymaking: fad or fasttrack?" (2017) 50 Policy Sci 367, doi: 10.1007/s11077-017-9293-1; Sarah Giest and Annemarie Samuels, "'For good measure': data gaps in a big data world" (2020) 53 Policy Sci 559, doi: 10.1007/s11077-020-09384-1; Nanna Bonde Thylstrup and others, "Politics of data reuse in machine learning systems: Theorizing reuse entanglements" (2022) 9 Big Data & Society 20539517221139785, doi: 10.1177/20539517221139785.

[7] Walter Castelnovo and Maddalena Sorrentino, "The Nodality Disconnect of Data-Driven Government" (2021) Administration & Society, doi: 10.1177/0095399721998689.

[8] Julia Black, *Rules and Regulators* (Oxford University Press 1997).

[9] Eric Windholz, *Governing Through Regulation: Public Policy, Regulation and the Law* (Taylor & Francis 2017).

[10] Sophie Weerts, "Le droit" in Primin Bundi Nils Sogel, Tobias Mettler, Sophie Weerts (ed), Comprendre et concevoir l'administration publique, le modèle IDHEAP (EPFL Press 2023).

[11] Giest (n 5); Sara Hughes, Sarah Giest and Laura Tozer, "Accountability and data-driven urban climate governance" (2020) 10 Nature Climate Change 1085, doi: 10.1038/s41558-020-00953-z; Valle-Cruz (n 5).

[12] Roger Brownsword, "Law disrupted, law re-imagined, law re-invented" (2019) 2019 Technology and Regulation 10; Lukas Lorenz, Judith van Erp and Albert Meijer, "Machine-learning algorithms in regulatory practice: Nine organisational challenges for regulatory agencies" (2022) 2022 Technology and Regulation 1, doi: 10.26116/techreg.2022.001; Lena Ulbricht and Karen Yeung, "Algorithmic regulation: A maturing concept for investigating regulation of and through algorithms" (2022) 16 Regulation & Governance 3, doi: 10.1111/rego.12437; Karen Yeung, "Algorithmic regulation: A critical interrogation" (2018) 12 Regulation & Governance 505, doi: 10.1111/rego.12158; Stephen L Roberts, "Big Data, algorithmic governmentality and the regulation of pandemic risk" (2019) 10 European Journal of Risk Regulation 94.

the legislative and regulatory processes is still missing. This paper aims to fill this gap and proposes the following research question: How would data-driven policy impact the regulatory process? To address this question, we delve into the work of the Organisation for Economic Co-operation and Development (OECD) to map how this institution envisions the integration of data-driven technologies into the public policy cycle. Additionally, drawing on Critical Data Studies (CDS), we aim to comprehend the potential effects of this approach on legislative and regulatory processes.

The work of the OECD in this field is highly relevant, as it stands at the forefront of discussions on harnessing data and data-driven innovation to revolutionise the public sector.[13] Indeed, the institution leverages its know-how and institutional capacity in policymaking to prepare documents and reports to guide national governments in transforming governments[14]. It conceives at conceptual and ideational levels what governments will implement with different engagement and success rates.[15] In this sense, this new wave of transformation in the policy cycle aligns with previous governmental changes driven by the adoption of information and technologies in the inner functioning of governments.[16] It follows the ideas of mGovernment, e-government, and digital government, representing distinct stages of adopting technology within government functions.[17] Over time, the OECD has actively engaged with such concepts and encouraged governments to take steps to complete this transformation. Finally, the OECD has proposed a new design of policymaking and regulatory processes, one that is anchored on new technological breakthroughs of data-driven innovation.[18]

The OECD's perspective on the future of policy-making and regulatory processes holds considerable significance and influence over governments[19]. The in-depth study of the OECD's work on the matter enables a comprehensive view of how emergent technologies can affect the regulatory process. By mapping how governments are being guided to undertake the digital transformation in what concerns regulation and policymaking, this paper proposes a reflection on the consequences of this transformation. It takes the perspective of CDS and regulation theory to analyse and question the OECD's vision for the future of policymaking and regulatory processes. The paper continues as follows. The next sections present the theoretical background, with insights from regulation and public policy theory (II), and critical data studies (III). The following section (IV) presents the analytical frame and the methodological approach. Section 5 showcases the work of the OECD and its view of the future of regulation and policy-making process. The OECD vision is discussed in section VI, and the conclusions are presented in section VII.

[13] Alice Chessé and Amy Verdun, "International political economy" in Fabrizio De Francesco and Claudio M. Radaelli (eds), *The Elgar Companion to the OECD* (Edward Elgar Publishing 2023) 10.4337/9781800886872; Clarissa Valli Buttow and Sophie Weerts, "Open Government Data: the OECD's Swiss Army Knife in the transformation of government" (2022) 14 Policy & Internet 219, doi: 10.1002/poi3.275.

[14] Matthieu Leimgruber and Matthias Schmelzer, "The historical transformations of the OECD" in Fabrizio De Francesco and Claudio M. Radaelli (eds), *The Elgar Companion to the OECD* (Edward Elgar Publishing 2023), doi: 10.4337/9781800886872.00008.

[15] Chessé and Verdun (n 13).

[16] Barbara Ubaldi and others, "*State of the art in the use of emerging technologies in the public sector*"(2019) Documents de travail de l'OCDE sur la gouvernance publique, n° 31, Éditions OCDE, Paris doi: 10.1787/932780bc-en.

[17] Tobias Mettler, "Transformation digitale" in Nils Soguel and others (eds), *Comprendre et concevoir l'administration publique : Le modèle IDHEAP* (Première édtion edn, EPFL Press 2023).

[18] Berryhill and others (n 1).

[19] Fabrizio De Francesco and Claudio M. Radaelli, *The Elgar Companion to the OECD* (Edward Elgar Publishing 2023).

II. Re-Contextualising data-driven public policy and rulemaking process

To better understand the possible effects of the new features of the policy cycle on the legislative and regulatory processes, it is important to disentangle the concepts and assert the connection between them.

First of all, public policy can be conceptualised as the authoritative governmental intervention in a given domain and territory designed to achieve the government's objectives.[20] Some authors add the nuance that, although adopted by governments, public policies emanate from society with the participation of various actors.[21] This intervention takes the form of different acts, such as plans, strategies, decisions, regulations and laws.[22] In this context, public policies are anchored in law and, at the same time, are also the frame in which new legislation and regulation come into being.[23] So, if law is an important shaper of policies, policies are the broad concept of governmental action that can cover legal acts and regulations.[24] In this way, new laws can be enacted as part of a new public policy. With this intricate relation in mind, we can understand that the policy-making process gives rise to new public policies, laws and regulations.

Several theories explain how public policy and regulations come to being.[25] Within the scope of this paper, it is crucial to recognise that the government's action relies on plausible hypotheses carefully formulated to achieve the government's objectives and desired outcomes, employing a rational method.[26] This rational method is not isolated from the social and political context, and therefore, while remaining rational, it is influenced and constrained by external and internal factors and actors. This reality introduces complexities and potential frictions with straightforward rationality. Nonetheless, the idea of a rational policymaker and a rational legislator is essential to justify its democratic legitimacy.[27] Within this framework, the quality of a regulation or a policy can be measured by its effectiveness in achieving and realising the public interest with the fewest resources and the lightest possible burden to private actors.[28]

Striving for effective and efficient policies, literature and international institutions have appealed to the so-called evidence-based public policy approach.[29] This movement aims to anchor regulation and policy processes into "theoretical informed empirical analysis" to achieve better results and avoid regulatory failure.[30] The idea of anchoring policy and regulatory processes in state-of-the-art knowledge can appear self-evident. Still, critics have been addressed to this approach because of the controversies regarding what can or cannot be considered as evidence and what knowledge is good enough to justify policy and regulatory decisions.[31] In this sense, questions have been raised concerning, for example, the admissibility of statistical evidence in the policy processes

[20] Windholz (n 9).

[21] Michael Howlett and Ishani Mukherjee, *Handbook of Policy Formulation* (Edward Elgar Publishing 2017).

[22] Jean-Claude Thoenig, "Politique publique" in *Dictionnaire des politiques publiques* (Presses de Sciences Po 2010) 10.3917/scpo.bouss.2010.01.0420.

[23] Weerts (n 10).

[24] Windholz (n 9).

[25] Robert Baldwin and others, "Explaining Regulation" in *Understanding Regulation: Theory, Strategy, and Practice* (Oxford Academic, Oxford University Press 2011).

[26] Windholz (n 9).

[27] Alexandre Flückiger, *(Re)faire la loi : traité de légistique à l'ère du droit souple* (Stämpfli 2019).

[28] Windholz (n 9).

[29] Holger Strassheim, "Trends towards Evidence-Based Policy Formulation" in Michael Howlett and Ishani Mukherjee (eds), *Handbook of Policy Formulation* (Edward Elgar Publishing 2017), doi: 10.4337/9781784719326.00042.

[30] Michael Howlett, "*Policy analytical capacity and evidence-based policy-making: Lessons from Canada*" (2009) 52 Canadian Public Administration 153, doi: 10.1111/j.1754-7121.2009.00070_1.x.

[31] Greg Marston and Rob Watts, "*Tampering with the evidence: a critical appraisal of evidence-based policy-making*" (2003) 3 The drawing board: An Australian review of public affairs 143.

but the dismissal of qualitative evidence. Others have debated the weight that should be given to evidence in the decision-making process. In one hand, some criticised the evidence-based approach for over considering the experts knowledge, noting the re-introduction of the idea of a technocracy as a way of imposing apolitical truths and thus denying the democratic debate.[32] Other authors point out the fact that politicians do not give due attention to evidence or select the most convenient evidence according to political goals and opportunities.[33] Finally, the approach has also been criticised for foreclosing the debate for many organisations that could not mobilise the resources to provide evidence and thus participate in the debate.[34] The overall balance of the evidence-based approach to public policy is up for debate, with several critics addressing its ability to improve the policy and regulatory processes.[35]

The data-driven public policy appears as the continuity of the evidence-based approach and, simultaneously, as a response to its limitations.[36] The idea of using data to improve policy and regulatory processes gained track with the improved possibilities of capturing ever more diverse and voluminous data and the emergence of innovative data processing techniques. This combination is perceived as enabling a better understanding of social problems and even the anticipation of the populations' needs.[37] Such an approach is perceived as responding to the question of what evidence is deemed good enough because data analytics can consider a wide variety of data (for example, it could combine social media data, data input from public services, and statistical data).[38] On the other hand, the debate on how to weigh the evidence is addressed by the idea that algorithms can properly evaluate and weigh which data should be considered in a given problem and how.[39] Concerning the actors' ability to participate in the public debate, the data-driven approach is deemed more democratic because it relies on multiple data inputs that, in theory, are open to the participation of all.[40] Additionally, collecting real-time data would enable a continuous feedback loop, enabling a somehow automatic re-assessment of the policy and, therefore, a responsive and adaptative regulatory approach.[41]

To which degree and how data-driven artificial intelligence will be embraced by governments remain a topic of debate from an empirical and normative point of view.[42] Indeed, governments' actual use of AI today remains an ongoing process with different levels of implementation. Additionally, part of the literature has pointed out the possible

[32] Marston and Watts (n 31).

[33] Strassheim (n 29).

[34] Howlett (n 30).

[35] Kati Rantala, Noora Alasuutari and Jaakko Kuorikoski, "*The logic of regulatory impact assessment: From evidence to evidential reasoning*" (2024) n/a Regulation & Governance, doi: 10.1111/rego.12542; Strassheim (n 28).

[36] Giest (n 6); Sofia Ranchordas and Abram Klop, "Data-Driven Regulation and Governance in Smart Cities" in *Research Handbook in Data Science and Law* (Edward Elgar Publishing 2018); Pedro Rubim Borges Fortes, Pablo Marcello Baquero and David Restrepo Amariles, "Artificial intelligence risks and algorithmic regulation" (2022) 13 European Journal of Risk Regulation, doi: 357 10.1017/err.2022.14.

[37] Margetts and Dorobantu (n 5).

[38] Somayeh Labafi and others, "Using an evidence-based approach for policy-making based on big data analysis and applying detection techniques on Twitter" (2022) 6 Big Data and Cognitive Computing 160.

[39] Labafi and others (n 38); Stefano Maffei, Francesco Leoni and Beatrice Villari, "Data-driven anticipatory governance. Emerging scenarios in data for policy practices" (2020) 3 Policy Design and Practice 123, doi: 10.1080/25741292.2020.1763896; R. Maule, G. Schacher and S. Gallup, "Knowledge management for the analysis of complex experimentation" (2002) 12 Internet Research 427, doi: 10.1108/10662240210447173.

[40] Cavaliere P and Romeo G, "From poisons to antidotes: Algorithms as democracy boosters" (2022) 13 European Journal of Risk Regulation, doi: 421 10.1017/err.2021.57.

[41] Benjamin Welby, "The impact of digital government on citizen well-being" (2019), doi: 10.1787/24bac82f-en.

[42] Anneke Zuiderwijk, Yu-Che Chen and Fadi Salem, "Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda" (2021) 38 Government Information Quarterly 101577, doi: 10.1016/j.giq.2021.101577.

limitations of this approach, noting that the imaginary and the rhetoric of data (re)use is far from the technicalities of real applications and that there are substantial technical challenges to the implementation of these visions.[43] Beyond the challenges of transforming this vision in reality, it is important to question the effects that such transformation could bring besides the supposed gains in efficiency and responsiveness. The literature on law and public policy has explored some of the possible effects of this new approach, especially the replication of bias and the intensification of discriminatory and surveillance practices.[44] Additionally, how the use of data-driven technologies as algorithmic governance systems influence the development and application of rules has been studied from a political sciences point of view[45]. However, this literature does not explore what this could mean to the legislative and regulatory process from the point of view of the formulation and adoption of rules. The CDS literature clarifies some of the limitations of the data-driven approach in general, and this can illuminate the theorisation of the effects that such an approach could have on the legislative and regulatory process.

III. Critical Data Studies and the limitations of data-driven approaches

CDS stands for an emergent interdisciplinary literature that studies the implications of the broad use of data and data-driven technologies in society. The CDS literature critically assesses how data is collected and interpreted within data assemblages and the ethical, social and political impact of its uses[46]. It finds its roots in the Sciences and Technology Studies (STS). Within the STS, one of the central tenets is the acknowledgement that as innovation becomes more pervasive in our lives, it exercises a transforming force in society, mutating the context in which it is deployed.[47] This transforming force of innovation is said to act as a colonising force, imposing specific ways of functioning through power struggles.[48]

In the following, we highlight certain concepts derived from the literature, offering insights pertinent to comprehending the potential transformation of regulatory and legislative processes through the pervasive integration of data-driven technologies in the policy cycle. In doing so, we do not have the ambition to cover all the possible contributions of CDS to the regulation theory. Instead, we present the foundational theoretical propositions of CDS at a more abstract level, providing a basis for reflection regarding the overarching impacts of this approach. Scholars in regulation theory can consider these propositions as a starting point for discussions on the future trajectory of regulation.

[43] Sabina Leonelli, "Data Science in Times of Pan(dem)ic" (2021) 3 Harvard Data Science Review, doi: 10.1162/99608f92.fbb1bdd6; Thylstrup and others (n 6).

[44] Jane Battersby, "Data Gaps and the Politics of Data: Generating Appropriate Data for Food System Assessment in Cape Town, South Africa" in Alison Blay-Palmer and others (eds), *Sustainable Food System Assessment* (Routledge 2019); Jonas Lerman, "Big data and its exclusions" (2013) 66 Stan L Rev Online 55; Linnet Taylor and Ralph Schroeder, "Is bigger better? The emergence of big data as a tool for international development policy" (2015) 80 GeoJournal 503, doi: 10.1007/s10708-014-9603-5; Derave C, Genicot N and Hetmanska N, "The Risks of Trustworthy Artificial Intelligence: The Case of the European Travel Information and Authorisation System" (2022) 13 European Journal of Risk Regulation 389, 10.1017/err.2022.5; Roberts (n 12).

[45] Daria Gritsenko and Matthew Wood, "Algorithmic governance: A modes of governance approach" (2022) 16 Regulation & Governance 45 doi:https://doi.org/10.1111/rego.12367.

[46] Robert Kitchin and Tracey Lauriault, "Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work" in Thatcher j, Eckert J and Shears A (eds), *Thinking Big Data in Geography, New Regimes, New Research* (University of Nebraska 2014).

[47] Bruno Latour, "Give Me a Laboratory and I Will Raise the World"(1983) Science Observed: Perspectives on the Social Study of Science 141.

[48] Sebastian Pfotenhauer and others, "The Politics of Scaling"(2022) 52 Social Studies of Science 3, doi: 10.1177/03063127211048945.

1. The supposed neutrality of data and the open-endedness of its uses

The premise of data reuse is underlying the promotion of data-driven innovation. The encouragement of data-driven technologies is centred on the idea that, if appropriately managed, data can be easily and readily reused. In this context, data is considered an autonomous entity that can easily be extracted from its original context and deployed elsewhere.[49] In this way, data collected to inform economic policies could be easily transferable to other areas of public action, such as education or health. The CDS literature points out several problems resulting from this premise. First is the assumption that data is neutral, reliable and unambiguous.[50] This assumption is not necessarily true, all the most because any process of data collection is necessarily selective and exclusionary.[51] Data is always an interpretation of reality and not the reality itself.[52] This interpretation is highly dependent on the context of the collection and the primary ends of this collection.[53] In this sense, the CDS literature refers to data as being always contextual and relational.[54]

Data, being always an interpretation of reality and, therefore, a version of it, poses problems concerning what has been called hermeneutical injustice.[55] El Sayed and Prainsack explain that using data in the policy-making process can be problematic because the reliance on existing and available data could narrow the possibilities of defining problems and hamper the epistemic definition of the problem and its possible solutions.[56] In this sense, the data production, collection, and selection processes are always political, even if they operate far away from political scrutiny.[57] In this sense, authors argue for an epistemic justice that promotes a fair and equitable way of defining collectively accessed problems.[58]

As an answer to the data being an interpretation of reality, some have proposed the collection of even more data. However, the search for the comprehensiveness of data denies the fact that no matter how much data one could collect, data remains an interpretation, and there are limited versions of reality interpretation.[59] The illusion or mirage of datafying everything is a rhetorical artifice to justify and legitimise data collection that is ever far from accountability, leading to what the literature has pointed out as increased surveillance.[60] In other words, if the only response to the failure of the data is to produce or collect even more data, this would lead inexorably to increased surveillance.[61]

Furthermore, this vision of data relies on the idea that data could be an autonomous entity separate from the algorithms that use them and the humans that manage its collection and use.[62] This idea is contested in the CDS domain, which explains that data is entangled with algorithms and that they are ontologically and intrinsically

[49] Thylstrup and others (n 6).

[50] Kitchin and Lauriault (n 46); Thylstrup and others (n 6).

[51] Paola Ricaurte, "Data Epistemologies, The Coloniality of Power, and Resistance" (2019) 20 Television & New Media 350 doi: 10.1177/1527476419831640.

[52] Leonelli (n 43); Nanna Bonde Thylstrup, "Data Out of Place: Toxic Traces and the Politics Of Recycling"(2019) 6 Big Data & Society 2053951719875479, doi: 10.1177/2053951719875479.

[53] Anne Beaulieu and Sabina Leonelli, *Data and Society: A Critical Introduction* (SAGE Publications Sage CA 2022).

[54] Beaulieu and Leonelli (n 53).

[55] Beaulieu and Leonelli (n 53).

[56] Seliem El-Sayed and Barbara Prainsack, "Blue Chips and White Collars: Whose Data Science Is It?" (2021) 3 Harvard Data Science Review, doi: 10.1162/99608f92.ea29ee5a.

[57] Beaulieu and Leonelli (n 53).

[58] El-Sayed and Prainsack (n 56).

[59] Ricaurte (n 51)

[60] Jose van Dijck, "Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology" (2014) 12 Surveillance Studies Networ 197, doi:10.24908/ss.v12i2.4776; Shoshana Zuboff, *The age of surveillance capitalism: the fight for a human future at the new frontier of power* (PublicAffairs 2019).

[61] Beaulieu and Leonelli (n 53).

[62] Adam Burke, "Occluded Algorithms" (2019) 6 Big Data & Society 2053951719858743 doi: 10.1177/2053951719858743.

co-constructed.[63] In this sense, Thylstrup and colleagues explain "data entities as phenomena that emerge within, and are contingent upon, the techno-scientific apparatus in which they appear."[64] Therefore, it would be an illusion to think about algorithms that could be stripped of data biases if trained in biased data sets. Also, and at a more abstract level, because algorithms are ontologically linked with their training data set, the political features of this data set would be reproduced even if the original data set is repaired or replaced or if the algorithm is applied to a different set of data or problem.[65]

Adding a different layer to the discussion, Bellanova and de Goede study how data-driven solutions are politically and juridically established and how they deflect and influence imaginaries and political discussions as they become part of the infrastructure of public action (2022). With that, they show the mutually constitutive features of political and technological choices.[66]

2. *Future lock-ins and foreclosed futures*

Most of the envisioned potential of data-driven technologies relies on their predictive potential. Indeed, one of the most celebrated capacities of data-powered artificial intelligence is the capacity to detect data patterns and, by doing so, predict possible outcomes.[67] Beaulieu & Leonelli argue that one of the features of a datafied society is the fact that progressively, future social action is determined by the consultation to data sources and intermediated by sophisticated and complex infrastructures that are ever more distant from social and public scrutiny.[68]

Through data analytics, the past determines the options for future action.[69] Indeed, data-driven technologies use techniques such as machine learning to learn patterns in data sets and arrive at probabilistic solutions. When the possibilities of future action are "calculated" based on data, which are always partial and incomplete, the possibilities of the future are also restrained. Data, therefore, archives the future.

Moreover, the data analytics "outputs" are the single result of deep probabilistic calculations that, at each layer of calculus, roll out a multiplicity of possibilities that are considered non-existent in the final result.[70] At each calculus level, multiple future pathways are closed in favour of the most probabilistic outcome.[71] In this sense, Amoore explains that "with contemporary machine learning algorithms, doubt becomes transformed into a malleable arrangement of weighted probabilities. Though this arrangement of probabilities contains within it a multiplicity of doubts in the model, the algorithm nonetheless condenses this multiplicity to a single output. A decision is placed beyond doubt."[72] The inductive logic of data-driven technologies hides its inner functioning, which remains opaque even to programmers. Consequently, it is hard to fight back, contest, or doubt the outcome, given that we can never fully assess how the

[63] Louise Amoore, *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others* (Duke University Press 2020).
[64] Thylstrup and others (n 6).
[65] Thylstrup and others (n 6).
[66] Rocco Bellanova, Helena Carrapico and Denis Duez, "Digital/sovereignty and European security integration: an introduction" (2022) 31 European Security 337 10.1080/09662839.2022.2101887.
[67] Helene Friis Ratner and Kasper Elmholdt, "Algorithmic constructions of risk: Anticipating uncertain futures in child protection services" (2023) 10 Big Data & Society 20539517231186120 doi: 10.1177/20539517231186120.
[68] Beaulieu and Leonelli (n 53).
[69] Thylstrup and others (n 6).
[70] Amoore (n 63).
[71] Ratner & Elmholdt (n 67).
[72] Amoore (n 63), p. 134.

algorithm finds its output.[73] As result, it becomes ever more difficult to reopen the foreclosed future pathways. Alternative visions of the future became difficult to even imagine.[74] Arriving at a similar conclusion and highlighting the consequences of this logic, Burrell states, "When an algorithmic logic is deployed to bring order to the world, it colonises the future by its relentless predictive logic. Alternative political claims are foreclosed."'[75] These perspectives shed light on how specific technologies and approaches shape the present, constrain and influence the potential pathways available for future societal developments.

Finally, it is important to mention that both perspectives described here (the problem of data neutrality and future lock-ins) are relevant when considering data-driven technologies as automated decision-making systems and a decision-aid mechanism. Indeed, even when the technologies are not the ones making the call on decisions, the fact that they suggest a path or provide insights or information for the human in charge of the decision already alters human behaviour.[76]

All in all, to understand how these important features of data-driven technologies can impact the regulatory process, and since we are in the early stages of implementing those technologies, the upcoming section will focus on the OECD's ideational work. By examining how the OECD envisions the use of data in policy and regulation, we can gain valuable insights into what might happen as a result of this transformation.

IV. Analytical frame and methodology

As stated previously, the formulation and execution of public policy typically follow a rational path, where problems are diagnosed, and solutions are proposed; political discussions shape the adopted solutions, which are subject to different disputes of interests and values.[77] To map how the envisioned view of data-driven policy and regulatory processes is being conceived, the paper analyses the work of the OECD through the lens of the policy cycle framework. The scholarship has proposed different formulations for the policy cycle stages, for this analysis, we deploy the cycle as proposed by Windholz. Windholz breaks down the policy cycle into separate steps or stages of a circular sequence that begins with setting out the problem and understanding its nuances (defining), goes on to designing solutions (designing), and enabling decisions (deciding), implementing these decisions (implementing) and finally evaluating the impact of the new regulatory and policy measures (evaluating).[78] Different actors are involved in the process's different phases or stages, and each stage gives rise to different political disputes and discussions.[79] While the policy cycle is a useful framework "to better understand the regulatory endeavour",[80] the adoption of policies in real life often takes a more tortuous path than the

[73] Alexander Campolo and Katia Schwerzmann, "From rules to examples: Machine learning's type of authority" (2023) 10 Big Data & Society 20539517231188725 doi: 10.1177/20539517231188725.

[74] Amoore (n 63).

[75] Jenna Burrell, "the ethics of uncertainty – Louise Amoore, Cloud Ethics: Algorithms and the attributes of ourselves and others (Durham, NC, Duke University Press, 2020, 232 p.)" (2020) 61 European Journal of Sociology/Archives Européennes de Sociologie 409, doi: 10.1017/S0003975620000193, 410.

[76] David Almog and others, "AI Oversight and Human Mistakes: Evidence from Centre Court" (2024) abs/2401.16754, ArXivdoi: 10.48550/arXiv.2401.16754; Robert Holton and Ross Boyd, "Where are the people? What are they doing? Why are they doing it?'(Mindell) Situating artificial intelligence within a socio-technical framework" (2021) 57 Journal of Sociology 179, doi: 10.1177/1440783319873046; Maia Jacobs and others, "How machine-learning recommendations influence clinician treatment selections: the example of antidepressant selection" (2021) 11 Translational Psychiatry 108, doi: 10.1038/s41398-021-01224-x.

[77] Windholz (n 9).

[78] Windholz (n 9).

[79] Thoenig (n 22).

[80] Windholz (n 9), p.85.

linear process proposed in the framework. In this sense, different phases can coexist, overlap, come back and forward or even be omitted in the process of policymaking.[81] Nonetheless, the cycle is preconised by the institutions working on better regulation approaches,[82] and the literature attests to the proliferation of the policy cycle as an analytical frame.[83]

Indeed, the policy cycle approach is deemed to insert rationality into the regulatory process, collaborating with what has been called, in the regulation theory, the proceduralisation of regulation.[84] Through the proceduratlisation of regulation, the legitimacy and efficacy of rules are enhanced by adopting procedures that ensure participation and, consequently, adherence to the regulation.[85] The policy cycle framework also enables highlighting how political ideas are conveyed, discussed and constructed throughout the adaptation and implementation of a policy.[86] This perspective also supports evidence that regulatory processes are communicative processes, and through this communicative process, legitimacy is reinforced and adherence to rules is constructed[87].

Through the analysis of the policy documents produced by the OECD, it is possible to notice that the uses of data-driven technologies are multiple and permeate all the stages of the regulatory enterprise. However, making this view real remains a challenge. To overcome this challenge, policy papers provide guidance and describe the future of data-driven technologies within the government.

To construct this analysis, the OECD's database/library was consulted. The "working and policy papers" category was selected within the library. Pertinent subcategories of these papers were scanned, and a total of fourteen subcategories were selected. Relevant documents published between January 2019 and September 2023 were chosen from these subcategories. The selection process focused on papers that, either in their titles or short descriptions, addressed the link between policy, public services and new technologies. The initial collection resulted in a list of 68 documents, which were thoroughly reviewed. In the second selection phase, this collection was narrowed down to twenty-two documents for an in-depth analysis. Table 1 presents the final list of the analysed documents.

Through content analysis and coding of these documents,[88] the study reorganises the technological interventions foreseen by the OECD and locates them within different steps of the policy cycle. The analysis codebook encompassed two axes of examination. The first axis pertains to policy-cycle steps: defining, designing, deciding, implementing, and evaluating. The second axis addresses the salience of CDS-sensitive content. For this second axis of analysis, the code book was defined based on the study of the CDS literature, and the established codes convey the main concepts that enable the discussion on the neutrality of data, its reusability and future lock-ins. The codes used in this second axis were: (re)usability of data, anchoring effect, co-creation of tech solutions, foreclosing the future, lack of transparency/obscured political decisions, and epistemological problems.

[81] Thoenig (n 22).

[82] OECD, "*Better Regulation Practices across the European Union 2022*" (2022) Éditions OCDE, Parishttps://doi.org/10.1787/6e4b095d-en.

[83] Thoenig (n 22).

[84] Julia Black, "Proceduralizing Regulation: Part I" (2000) 20 Oxford Journal of Legal Studies 597.

[85] Black (n 84).

[86] Julia Black, "Regulatory Conversations" (2002) 29 Journal of Law and Society 163 doi: 10.1111/1467-6478.00215.

[87] Robert Baldwin, Martin Cave and Martin Lodge, *Understanding Regulation: Theory, Strategy, and Practice* (2 ed, Oxford Scholarship Online 2011).

[88] Lindsay Prior, "Content Analysis" in Patricia Leavy (ed), *The Oxford Handbook of Qualitative Research* (Oxford Academic, Oxford University Press 2020) 10.1093/oxfordhb/9780190847388.013.25 <https://doi.org/10.1093/oxfordhb/9780190847388.013.25> accessed 19/04/2024.

Table 1. List of documents

Final list of analysed documents	
Number	Reference
1.	OECD (2020), "Going Digital integrated policy framework," OECD Digital Economy Papers, No. 292, OECD Publishing, Paris, https://doi.org/10.1787/dc930adc-en.
2.	OECD (2019), "Using digital technologies to improve the design and enforcement of public policies," OECD Digital Economy Papers, No. 274, OECD Publishing, Paris, https://doi.org/10.1787/99b9ba70-en.
3.	OECD (2023), "Regulatory sandboxes in artificial intelligence," OECD Digital Economy Papers, No. 356, OECD Publishing, Paris, https://doi.org/10.1787/8f80a0e6-en.
4.	OECD (2022), "Responding to societal challenges with data: Access, sharing, stewardship and control," OECD Digital Economy Papers, No. 342, OECD Publishing, Paris, https://doi.org/10.1787/2182ce9f-en.
5.	OECD (2022), "The OECD Going Digital Measurement Roadmap," OECD Digital Economy Papers, No. 328, OECD Publishing, Paris, https://doi.org/10.1787/bd10100f-en.
6.	OECD. (2021). Achieving cross-border government innovation: Surfacing insights and experimenting across borders. doi:doi:https://doi.org/10.1787/3535f0ea-en
7.	OECD. (2021). Achieving cross-border government innovation: Governing cross-border challenges. doi:doi:https://doi.org/10.1787/ddd07e3b-en
8.	OECD. (2020). The OECD Digital Government Policy Framework. Six dimensions of a Digital Government doi:doi:https://doi.org/10.1787/f64fed2a-en
9.	OECD. (2022). Achieving cross-border government innovation: Delivering and enabling impactful solutions. doi:doi:https://doi.org/10.1787/cb6dd61b-en
10.	OECD. (2022). OECD Good Practice Principles for Public Service Design and Delivery in the Digital Age. doi:doi:https://doi.org/10.1787/2ade500b-en.
11.	Welby, B., & Tan, E. H. Y. (2022). Designing and delivering public services in the digital age. doi: doi:https://doi.org/10.1787/e056ef99-en
12.	Monteiro, B. and R. Dal Borgo (2023), "Supporting decision making with strategic foresight: An emerging framework for proactive and prospective governments," OECD Working Papers on Public Governance, No. 63, OECD Publishing, Paris, https://doi.org/10.1787/1d78c791-en.
13.	Kaur, M, et al. (2022), "Innovative capacity of governments: A systemic framework," OECD Working Papers on Public Governance, No. 51, OECD Publishing, Paris, https://doi.org/10.1787/52389006-en.
14.	OECD (2021), "The OECD Framework for digital talent and skills in the public sector," OECD Working Papers on Public Governance, No. 45, OECD Publishing, Paris, https://doi.org/10.1787/4e7c3f58-en.
15.	Tõnurist, P. and A Hanson (2020), "Anticipatory innovation governance: Shaping the future through proactive policy making," OECD Working Papers on Public Governance, No. 44, OECD Publishing, Paris, https://doi.org/10.1787/cce14d80-en.
16.	Mohun, J. and A. Roberts (2020), "Cracking the code: Rulemaking for humans and machines," OECD Working Papers on Public Governance, No. 42, OECD Publishing, Paris, https://doi.org/10.1787/3afe6ba5-en.
17.	Ubaldi, B. and T. Okubo (2020), "OECD Digital Government Index (DGI): Methodology and 2019 results," OECD Working Papers on Public Governance, No. 41, OECD Publishing, Paris, https://doi.org/10.1787/b00142a4-en.
18.	Berryhill, J, et al. (2019), "Hello, World: Artificial intelligence and its use in the public sector," OECD Working Papers on Public Governance, No. 36, OECD Publishing, Paris, https://doi.org/10.1787/726fd39d-en.

(*Continued*)

Table 1. (*Continued*)

Final list of analysed documents	
Number	Reference
19.	van Ooijen, C, B. Ubaldi and B. Welby (2019), "A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance," OECD Working Papers on Public Governance, No. 33, OECD Publishing, Paris, https://doi.org/10.1787/09ab162c-en.
20.	Welby, B. (2019), "The impact of digital government on citizen well-being," OECD Working Papers on Public Governance, No. 32, OECD Publishing, Paris, https://doi.org/10.1787/24bac82f-en.
21.	Ubaldi, B, et al. (2019), "State of the art in the use of emerging technologies in the public sector," OECD Working Papers on Public Governance, No. 31, OECD Publishing, Paris, https://doi.org/10.1787/932780bc-en.
22.	Galindo, L, Perset, K. and Sheeka, F. (2021). An overview of national AI strategies and policies. doi:doi:https://doi.org/10.1787/c05140d9-en

Table 2. Codebook

Policy Cycle	Defining
	Designing
	Deciding
	Implementing
	Evaluating
CDS-sensitive content	(re)Usability of data
	Anchoring effect
	Co-creation of tech solutions
	Foreclosing the future
	Lack of transparency/obscured political decisions
	Epistemological problems

The analysis was carried out with the help of the software Atlas ti. Table 2 presents the codebook.

V. The place of new technologies into the policy and regulation-making process

The OECD's overarching goal is to introduce technological solutions to achieve efficiency and improve the quality of public services. In its discourse, there is a general incentive to adopt innovative approaches to the regulatory process, with different solutions being discussed in several instances. There are also caution alerts concerning the possible pitfalls of these approaches. However, they are always relativised in the face of the alleged greater beneficial potential of the innovative process. If there is no general step-by-step orientation concerning how this technological turn will be implemented, there is a discourse encouraging experimentation and exploratory measures (Document 18). In the

following sections, we decouple the different views of technology and their foreseen intervention in each stage of the policy and regulatory cycle.

1. Defining

The defining phase of the policy and regulatory cycle analyses the public problem at hand, tracing back its causes and evaluating its consequences. In this phase, with data-driven innovation, policymakers would understand the current situation better by using existing data from different contexts and sources. Data analytics would be a key resource to help assess the contours of problems, infer possible causes, identify affected actors, and measure the problem's consequences (Document 2). In this context, the "increased availability, quantity, complexity and production rate of data, together with revolutionary developments in data analytics" (Document 19) are highlighted as central elements, enabling a true transformation in the inner functioning of governments. Within this view, the tasks of government should be rethought to harvest the most of data analytics opportunities. This involves amplifying data collection and ameliorating data analytics throughout the government's activity. By undertaking these measures, governments can access a broader and more diverse range of data, enabling a more comprehensive evaluation of public problems. The definition phase is, in that way, rethought on the basis of available data, and efforts need to be undertaken to increase opportunities for data collection (Document 8).

Within the possibilities of using data analytics to define problems, a more ambitious view of policy design aims to build a proactive government in forecasting and foreseeing problems, not just diagnosing present issues. In this sense, for instance, the work of the OECD underlines that: "Data analytics can reveal social and economic trends and incite a transition from reactive to more proactive and forward-looking data-based policymaking and service delivery." (Document 19, p. 9).

2. Designing

In this phase, a range of possible solutions is envisioned and compared. Policymakers need to assess the potential of different policy paths and evaluate available tools to address the problem (which was defined in the previous phase). In this stage, forecast data analytics is presented as a tool enabling governments to understand better how different solutions could lead to different outcomes. In this sense: "This is not about predicting whether a policy will be effective, but about how countries experiment with particular approaches or consider a wider spectrum of intervention in order to shape future policy interventions, informed by the availability of data." (Document 19, p. 19). Additionally, data-driven technologies can be used to forecast how the target population would respond and comply with different regulatory measures: "machine learning (ML), algorithms can be used to predict the sub-sample of the potential target population that is most likely to be 'policy-compliant,' i.e. responsive to the intervention." (Document 2, p. 23).

At a more conceptual level, the OECD envisages bringing the design phase closer to the implementation phase. In this sense, there is a discourse that encourages the conception of technical tools for implementation already in the design phase. Indeed, in the view of the OECD, the policy decisions are "isolated from delivery realities, while the operational experience is contained within a further, separate silo." (Document 11, p. 9). This reality would hamper the achievement of the policy goals and "ultimately affect trust in government" (Document 11, p. 9). Therefore, there is an encouragement to adopt, at the design phase, approaches that use exploration, research and prototyping aiming to test diverse solutions and adopt the best fit (Documemt 2). This means developing technical solutions and thinking about data governance at an early stage (in the design phase) of the

policy cycle. In this sense: "When policy development takes place in close proximity to the design and delivery of a service, research findings and experimental, hypothesis-led interventions can be incorporated into the service itself and tested quickly." (Document 11, p. 11) This conceptualisation of implementation in a very early stage would enable the consideration of technical mechanisms as instruments of public action. In this sense, data-driven technologies become instruments of public action. For example, the OECD mentions the use of data-driven tools in the environmental policy: digital nudges to reduce energy consumption, direct consumers to greener choices, and sustainable online shopping.

3. Deciding

In this stage, the course of action is selected by the policymaker. In the OECD vision, the decision stage is also permeated by data-driven technology. Here, the data analytics tools are used to showcase important information and to improve political discussions on the issue: "Data analytics can support the process of debating policy options with stakeholders by providing detailed information about a policy problem in an accessible way. Methods such as interactive data visualisation help make conceptual issues real in ways that can be presented for different audiences and that provide specific entry points to engage stakeholders in contributing potential solutions." (Document 19, p. 20).

Data-driven technologies are also proposed as solutions to gather stakeholders' perspectives and provide decision-makers with a broad and comprehensive view of inputs of the participation process. "Besides facilitating traditional methods of stakeholder engagement, data analytics can help governments analyse the wealth of input coming from crowdsourcing initiatives. Automated analysis of data regarding citizens' opinions and/or behaviours through such channels as public blogs, social media or public consultations provides insight on the public reactions towards certain policy problems." (Document 19, p. 20). The decisions anchored in data-driven techno-solutions are portrayed as more fair and accurate than the ones using traditional or analogical decision-making (Document 18).

4. Implementing

The implementation phase consists of translating the political decisions of the previous phase into real-life measures. Here, data-driven technologies are conceived as mechanisms for monitoring performance, auditioning and automating decisions, and, lastly, automated policy implementation (rule-as-code).

Two possibilities are pointed out concerning the use of data-driven technologies as monitoring and auditioning tools. First, authorities could rely on data produced by society. The example mentioned is the adoption of mechanisms that allow citizens to inform and report situations such as road problems, suspicious behaviour and other everyday issues (Documnt 19, p. 21). Second, a more sophisticated data analysis would involve the use of data produced by society combined with the data produced throughout the service delivery, with the idea of creating an 'ongoing performance monitoring and management' (Document 19, p. 22). The idea of "ongoing performance monitoring" depends on collecting real-time data. Therefore, in this vision, the implementation phase should also be conceived as a key opportunity for data collection. In some cases, the collection of granular data would enable the implementation of tailored (or differentiated) policy, "in which marginal benefits and cost are more closely aligned." (Document 2, p.7). The tailored regulation is also called policy differentiation, and it could be imaged as differentiation in terms of time (for example, different tax regimes to different delays in tax-returns declarations), space (for instance, different benefits can target different geographical areas), or technology (for example, different rules for convention combustion and electric

vehicles). The policy differentiation that was previously very difficult to implement becomes feasible with the mediation of data-driven technologies.

Data-driven tools are also considered important tools for compliance and enforcement of regulations. The gathering of data would enable "improved monitoring and supervision and more effective enforcement of policies" (Document 2, p.6). These tools would help compliance surveillance by indicating where to deploy investigation efforts (Document 18, Document 2). The example comes from technologies used to help detect fraud in tax or social benefits claims (Document 21, p. 28). Ultimately, technologies could be used to assist and automate decision-making. Automating decisions is certainly one of the more controversial areas of application of data-driven technologies. The idea is that, on the basis of the datafication of previous decisions, it would be feasible to automate future decisions. However controversial, its applications are ever more present in public administration, and several examples are mentioned by the OECD (examples of application include access to higher education (Document 2), the process of application (Document 16), and access to public procurement (Document 22)).

Within the implementation phase, the idea of the Rule-as-Code (RaC) approach appears, in which the conception of the rule, or law, is embedded in the technological solution for its implementation. The approach is also called machine-ready legislation. Again, it is possible to see the idea of bringing together the design and the implementation phases: "By reducing the number of opportunities for misinterpretation between the designers and implementers of policy, RaC can deliver policy outcomes more true to their original intent" (Document 16, p. 40). The approach would also improve the policy or regulation monitoring: "By minimising the opportunity for misinterpretation, it will not only be easier to see if the rules are having the desired effect but also if or where any implementation issues with those rules may lie" (Document 16, p. 40). Although in this approach, the rule is encoded and, therefore, its automation is not a data-driven technology, there is the idea that this approach would enable data collection in real time. Consequently, the data collected by the automated rule would enter the ecosystem and enable the policy/regulation evaluation and, ultimately, its review. Finally, the approach would enable the in-real-time monitoring of the rule and its effectiveness: "In this way, humans will act in accordance with the rules embedded in digital infrastructure, even though the rules themselves may no longer be immediately apparent or visible" (Document 16, p. 20).

5. Evaluating

The last step in the regulatory policy process is evaluation. Evaluation means methodical assessment of all or parts of a regulatory policy with the aim of providing recommendations. Here, data-driven innovation should be used in a retrospective sense, helping policymakers and regulators understand the effect of regulations and policies and anticipating and forecasting possible changes, new needs or desirable improvements (Document 19). The evaluation of the efficiency of existing processes would rely on data analytics. The possibility of increased data collection and analysis in the implementation phase would enable an evaluation that is not distant in time but continuously occurring. In this sense, the proposed view would become an ongoing real-time process: "As a result, policy evaluation can turn into an open, ongoing process rather than an internal, snapshot moment. Data analytics can be especially powerful in identifying the adverse effects of a decision or a policy and being able to respond accordingly" (Document 19, p. 26).

Consequently, it would enable quick adjustments and a more responsive government. In the work of the OECD, the advantages of implementing a real-time continuous evaluation system are somewhat relativised: "The availability of real-time social and economic data entails the potential for frequent policy adjustment to address changing market

conditions, yet the continuous reformulation of policy may undermine stable, long term policy signals" (Document 2, p. 17). Nonetheless, the idea of data collection to measure policy efficiency is deemed to be a key area of innovation (Document 2).

VI. Discussion

1. The reliance on (past) data as a building block of the future

The primary instrument enabling this novel approach to the policy cycle is a reimagined understanding of data within the government. Realising this vision is anchored in fostering a new data culture that integrates data gathering and analytics throughout the policy cycle. The adoption of such an approach would boost government intelligence and analytical capacity. Data-driven technologies would improve evidence-based policy, enhance public service delivery, and promote a more responsive and adaptative government.

The centrality of data is based on the idea that data should be collected from the most diverse sources and as often as possible (even continuously, in some cases). Then, such data could be reused to provide information not only for the contextual policy from which it was collected but also beyond it, to feed information to multiple policy areas (Document 19). In the same way, the OECD preconises the collection of the so-called "longitudinal administrative data," which captures data from a long-term perspective and enables the vision of an evolving situation (Document 2). However, this vision of data collides with the limitations that CDS has pointed out as epistemological and technical limitations for data reuse. Indeed, as CDS literature has shown, data is not neutrally collected or harvested, but it is constructed. In the construction of data, its context is highly relevant. This makes travelling data from one policy context to another a delicate task to which careful thought must be applied. In this sense, the rhetoric of data reuse can be deceiving because it undervalues the difficulties of considering data as contextual and relational. This could be problematic throughout the policy cycle, but it is more prominent in the definition of the problem, the design of policy options, and the decision-making.

First, concerning the definition of problems and their causes and consequences, the reliance on data can narrow the possibility of understanding a problem based on the existing data. The OECD's discourse emphasises the need for increased data and analysis as a tool to better understand current social problems and even to anticipate future social challenges; in this way, data works as the instrument helping to define the exact contours, causes and effects of public problems. However, the CDS literature already clarified that data works not as windows that enable one to see and understand reality but as a lens that frames reality in a certain way. In this sense, if only the datafied factors and actors count as the reality, this could misrepresent the complexity and depth of the problem. In this sense, data should be considered carefully, and this limitation should be made explicit.

In a similar way, the reliance on data collection could frame and determine the choice of solutions, and solutions where the data collection is more easily achieved could be preferred, even if they are not necessarily the more efficient or even if such solutions would impose a great burden to citizen (including incursions in citizens' privacy). Indeed, there is an increasing appeal to consider data collection opportunities as a substantial factor when choosing political options for public action. Such an approach would result in a preference for datafiable solutions to the detriment of low-tech solutions even when the low-tech would be more reasonable. Moreover, the insistence on adopting policy and regulatory solutions that enable data collection could result in the increased surveillance of citizens.

At the level of decision-making, this could also generate problems. The overreliance on data could narrow political debate on essential issues. The rhetoric of data neutrality could

hide important political choices that are made way before the public discussion of the matter. In this sense, the political choices underlying data collection and analysis could be occluded from decision-makers and held far from the public debate. The limitation of the political debate has important implications at the regulatory level. As mentioned above, the regulatory process is a communicative process. The communicative process is where regulatory conversations occur, and such conversations are essential to build adherence to rules and reinforce the legitimacy of both rules and authorities. Therefore, a substantial change in this process that becomes mediated by data-driven technologies could undermine rules' legitimacy and complicate compliance.

2. *Closing the gap between different policy cycle steps*

The envisioned view of a data-driven policy process brings closer steps in the policy cycle that were previously conceived as separate stages. First, the idea of bringing together the design and the implementation steps is advanced. The solution would adopt a design that enables experimentation. In this experiment, different solutions would be tested, for example, with a small scope, and then, the more efficient solution would be scaled up. Shifting the implementation stage to occur before the decision stage, even if just during the testing phase, will have an impact on the regulatory process. To conduct the "test," a preliminary regulatory framework must be established (due to the principle of legality). This could be achieved, at least in theory, by the adoption of experimental regulations, a tool already used in various fields.[89] However, widespread adoption of such regulatory tools may complicate the regulatory environment and result in legal uncertainty.[90] Furthermore, together with the idea of testing different policy solutions, there is the idea that such solutions would be data-driven. This means that the political solution is thought of together with the technological tool that would implement it. Such an approach could be interesting from the managerial point of view but could represent challenges for the regulatory process that becomes anchored in a technical solution. In what measure would the technicities of the tool be incorporated and integrated into the law? It could lead to problems, changing some of the main features of law and the legal systems, for example, the fact that law is mediated by language. Such a mutation could change the nature of law and legal system.[91] Therefore, a cautious approach should be favoured.

Second, the envisioned view of a data-driven policy brings the evaluation and defining stages closer together. The in-real-time monitoring of the policy makes the output (the policy or regulatory results) ever closer to the input (the development of a new policy cycle). Indeed, the data-driven policy cycle is based on the continuous data collection and analysis, enabling ever more rapid and dynamic adjustments. This transformation can cause a real disruption of the regulatory cycle as we know it. In this sense, the work of OECD illustrates this view: "An agile, research-led approach emphasises a continual response to learning from the experience of users and applying those insights to improve outcomes." (Document 11, p. 11). In this sense, the OECD praises the adoption of open-ended approaches to public policies or an adaptative policymaking process.[92] If this gap, or temporal space, between evaluating the results and defining new political problem/solutions is suppressed or automated, this could severely diminish the space of public

[89] Sofia Ranchordas and Valeria Vinci, "Regulatory Sandboxes and Innovation-friendly Regulation: Between Collaboration and Capture" (2024) forthcoming Italian Journal of Public Law.

[90] Ittai Bar-Siman-Tov, "Temporary legislation, better regulation, and experimentalist governance: An empirical study" (2018) 12 Regulation & Governance 192 doi: 10.1111/rego.12148.

[91] Mireille Hildebrandt, Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology (Edward Elgar 2015).

[92] Piret Tõnurist and Angela Hanson, "Anticipatory innovation governance" (2020) Documents de travail de l'OCDE sur la gouvernance publique, n° 44, Éditions OCDE, Parisdoi: 10.1787/cce14d80-en.

debate on the issue at hand. A possible consequence of such a shift in the policy cycle for regulation is that the scope of legal rules and the power of legislative bodies could decline. Indeed, given that the policy solution would need to adapt fast, the rules need to be conceived as open-ended solutions. The regulators would then be responsible only for prescribing the general and abstract principles. The administrative bodies and the technical teams would be reinforced in their power to implement these general and abstract principles into concrete measures and policies that are constantly adapting. With this, we can foresee a decline in the classic democratic institutions. It could be that alternatives to the classical democratic institutions would gain terrain, such as citizens-councils or platforms where specific policies and regulations could be scrutinised, with the evaluation of concrete technological solutions.

3. *The anticipatory governance and foreclosing futures*

Another aspect of the transformation of the policy cycle by data-driven innovation is the use of technology to forecast trends in behaviour and foresee future behaviour as an object of regulation. In this sense, data-driven technologies give rise to possibilities previously inconceivable, such as taking stock of the past, having a glance at the future, and acting on it to improve the chances of the desirable path and to pre-empt undesirable outcomes. Implementing predictive policies has the potential for profound repercussions at both the individual and collective levels. Individually, it may impose constraints, limiting individuals' ability to break free from socio-economic lock-ins. Moreover, at the collective level, the elimination of diverse policy pathways could lead to undesirable outcomes, creating a scenario where potential policy alternatives are stripped away. The exact contours of how this could take place in real life are still being defined. However, this possibility is perhaps the most problematic from the point of view of legitimacy and democracy.

First, it is important not to understate the problem of data being only one interpretation of reality. Therefore, it always represents a limited view of the reality. In this sense, the future constructed over a limited version of the past (datafied past) will be limited too. The question of epistemic justice gains a dimension for the future. A conception of a public problem that does not consider all the relevant points of view (because some of them are not datafied or hard to translate into data) will result in an intervention in the future that reproduces this epistemic injustice. The archived future reproduces the injustices of the data it is based on. Second, by relying on data-driven technologies to provide responses on the most efficient policy path, other policy options are foreclosed on the basis of probabilistic calculations. The underlying calculations remain far from public scrutiny and are not the object of political discussion, even if there are political choices hidden in the calculus layers. As algorithms convey a unique output that is placed beyond doubt, all the doubts and uncertainties of its statistical analysis are no longer objects of political discussion. In this way, the possibilities of collective and societal developments that are the affordances of incertitude could be restrained by the apparent sureness of data-driven responses.

In this scenario, and from the point of view of regulation, this could result in a decrease in the democratic legitimacy of regulation. Indeed, regulatory legitimacy is constructed, among other things, through regulatory conversations and political debates that promote their adoption. If such debates are suppressed or significantly diminished by the intervention of technological fixes, this could have an undesirable effect on its perceived legitimacy, impacting adherence to the norm and overall compliance.

VII. Conclusion

Our analysis shows that the OECD's view of a digital government would embrace data-driven technologies in all policy cycle stages. Beyond mere technological interventions, these advancements would also represent a significant transformation in the conceptualisation of the policy cycle: bringing together stages previously thought of as separate stages of the rational path of governmental action and using data-driven technologies to intermediate the tasks of thinking about a public problem and acting on it. Thus, these technologies provide an avenue for reimagining the entire process, making it more agile and iterative. Such transformation would also impact the regulatory process and regulation. The consequences of this transformation, from the regulatory point of view, can be summarised in the following points:

- It could impact the way public problems are conceived and frame the possibilities of action.
- The readability of data may emerge as a decisive factor in steering public policy choices, potentially favouring solutions that are more easily adaptable to a datafied world, regardless of possible pitfalls.
- Widespread adoption of experimental regulation could become mainstream, which would increase the complexity of legal systems and potentially create legal uncertainty.
- A reinforcement of administrative and implementation bodies with the adoption of legal norms that are more principle driven. And a policy implementation that is adaptable and open-ended.
- Data-driven technologies can obscure political choices that are "hidden" in data collection and data analytics. This could hamper the legitimacy of regulations and impact adherence and compliance with the norms.

Such changes in the regulatory process and regulation can be explained by the CDS insights, which underline that innovations exert a transformative influence on the environment they operate within. It is only natural, therefore, that the policymaking process undergoes adaptations to accommodate these technologies and harness their potential. Having a grasp of these consequences is useful to policymakers and the literature that may be better equipped to evaluate the challenges and opportunities of this transformation. Awareness of these consequences can also be an opportunity to discuss further solutions and find alternative paths that could preserve public debate on our collective futures.

At this point, it is also important to point out that our work is based on policy documents produced by the OECD, in which the view of the data-driven policy cycle is conceived at a conceptual level. Such a work mentions a few practical examples currently being implemented by diverse governments worldwide. However, the realisation of such ideas in the real world remains a research topic on its own. Moreover, the use of CDS concepts has brought to light several issues that are highly relevant to the discussion on data-driven policy; however, the points explored in this paper are not exhaustive, and the intersection between these two literatures promises fruitful paths of research, offering a rich landscape for additional exploration and analysis.

Cite this article: C Valli Buttow (2025). "Data-Driven Policy Making and Its Impacts on Regulation: A Study of the OECD Vision in the Light of Data Critical Studies". *European Journal of Risk Regulation* **16**, 114–132. https://doi.org/10.1017/err.2024.73