Building a 'Green Data' Future: how a human-centric approach to data and nudges can help fight climate change

Mark Fenwick and Paulius Jurcys*

1. Introduction

In November 2022, the global population surpassed 8 billion people.[1] The rapid growth of the human species is a testament to the progress that humanity has achieved in science, technology and healthcare. Still, the fate of humanity depends on whether and how quickly we are able to solve the challenges of climate change and its effects. An ongoing and pressing issue for the global community of regulators and other policymakers is indeed how to collectively address the climate change and the carbon problem issues we currently face. What tools should be put in place to ensure we meet our carbon reduction goals? What technologies, regulations and processes are needed to help us immediately contain the CO_2 emissions and reach net zero carbon emission targets?

One of the major causes of climate change is human activity and, more specifically, the use of fossil fuels that cause the emission of carbon dioxide. CO_2 results from driving cars, generating electricity, building infrastructure and enjoying what we consider a better quality of life. One way of framing the task that lies ahead of us with regard to decarbonization is how to enable economic growth without creating harmful environmental impacts.

The authors

- Mark Fenwick is a professor of International Business Law, Graduate School of Law, Kyushu University, Fukuoka, Japan. Paulius Jurcys is a senior research fellow at Vilnius University Law Faculty (Vilnius, Lithuania) and a co-founder of Prifina Inc. (San Francisco, USA).

Abstract

- A defining feature of digital transformation is the sensorization of everything and the concomitant creation of vast amounts of data covering every aspect of social and economic life. It is unfortunate that the value of this vast data pool has not been fully leveraged in tackling climate change and other environmental challenges facing humanity.
- This article suggests that this failure is a product of the closed and siloed character of the modern data ecosystem in which some stakeholders deem the collected data as their proprietary asset and have little incentive to share it with those interested in creating new value from such data, including, for example, by developing innovative technologies relevant for tackling climate change.
- This article argues that a user-held data model, in which data are reconceptualized as the property of the users who generate it, might better unlock or activate the 'green value' of these data and, by doing so, facilitate the creation of innovative green technologies. This paper focuses on three settings where a user-centric approach to data might usefully contribute to achieving these environmental goals, namely, facilitating green lifestyles, fostering sustainability in green cities and building green innovation ecosystems. Based on the foregoing discussion, this article concludes by identifying some normative principles that might be useful to guide regulatory interventions in this context.

* Emails: mdf0911@gmail.com; pjurcys@gmail.com

1 Tom Crowfoot 'World Population Just Passed 8 Billion. Here's What It Means' WEF. Available at https://www.weforum.org/agenda/2022/11/world-population-passes-8-billion-what-you-need-to-know/ (accessed 12 January 2023).

doi:https://doi.org/10.1093/jiplp/jpad031
Advance Access Publication 11 April 2023

Investments in green and sustainable technologies reached an all-time high in 2021.[2] In reality, however, paring down CO_2 emissions has become difficult because of the high number of global stakeholders and conflicting interests. CO_2 emission has been especially challenging in the most polluting industries (ie steel manufacturing, construction, mobility and agriculture). Therefore, new approaches are being developed to address the carbon equation problem.[3] One such promising endeavour, for example, revolves around the idea of capturing CO_2 that is already in the atmosphere and safely storing it for thousands of years.[4] To advance carbon capture technologies, some of the largest technology and consulting companies in Silicon Valley (namely, Stripe, Meta, Alphabet, Google and McKinsey) have launched a 1 billion US dollar fund that aims to provide advance funding for companies working on carbon capture technologies to help build the market for carbon capture.[5]

Many governments and public and private organizations have adopted various environment, social and governance (ESG) programmes to achieve net zero carbon emission objectives. There are three main elements of such net zero programmes: (i) to measure carbon emissions, (ii) to reduce CO_2 emissions as much as possible and (iii) to do everything else that is possible under the given circumstances.[6] This said, ESG initiatives on a corporate level have been rather ineffective as differences in data and metrics often make meaningful comparisons difficult.[7]

The above-mentioned strategies all involve various sets of data that must be harnessed and utilized in achieving the carbon equation. Ironically and disappointingly, the potential 'green value' of these data has been largely ignored. Although the amount of data generated by various sensors continues to grow exponentially, only a small amount of such data is used in practice to address the climate change issues.

Academic research in IP has mostly focused on the analysis of how existing IP frameworks and different areas of IP contribute to the global efforts to fight climate change. For instance, academics have been exploring role intersection between patent rights and unlocking sustainable/green technologies,[8] the right to repair[9] or how fashion brands are adopting new recycled materials, anticompetitive implications,[10] as well as risks of using misleading terms to define products as environmentally friendly (so-called 'greenwashing').[11] In the past several years, IP academics and practitioners have increasingly discussed issues related to the relationship between data and IP rights, and most recently, the debate has focused on access to data.[12]

One of the major obstacles in fighting climate change and reaching ESG goals is that most of the data are locked away behind the walled gardens of device manufacturers and large technology companies.[13] Data generated by consumers interacting with apps and using sensorized devices are collected, stored and processed

2 Bloomberg reports that investments in building energy efficiency increased 16% last year to an all-time high of $237 billion. See Gautam Naik 'Global Emissions from Buildings, Construction Climb to Record Levels' Available at https://www.bloomberg.com/news/articles/2022-11-09/global-emissions-from-buildings-construction-climb-to-record-levels?leadSource=uverify%20wall (accessed 12 January 2023).

3 DS Schimel, 'The carbon equation' (1998) 393 *Nature* 208.

4 See Leslie Kaufman and Laura Miller Lombrana 'Six Climate Breakthroughs That Made 2022 a Step Toward Net Zero' Available at https://www.bloomberg.com/news/articles/2022-12-29/2022-has-been-a-great-year-for-climate-progress-six-reasons?srnd=premium&leadSource=uverify%20wall (accessed 12 January 2023).

5 Available at https://frontierclimate.com/ (accessed 12 January 2023).

6 S Smith and N Ransohoff, 'The Economics of Carbon Removal with Nan Ransohoff' Available at https://a16z.simplecast.com/episodes/the-economics-of-carbon-removal (accessed 12 January 2023).

7 M Taddeo, A Tsamados, J Cowls and L Floridi, 'Artificial intelligence and the climate emergency: Opportunities, challenges, and recommendations' (2021) 4 *One Earth* 777; HS Sætra 'The AI ESG Protocol: Evaluating and Disclosing the Environment, Social, and Governance Implications of Artificial Intelligence Capabilities, Assets, and Activities' Available at https://onlinelibrary.wiley.com/doi/10.1002/sd.2438 (2022) (exploring how data and AI relate to company ESG performance, accessed 12 January 2023); AL Stein, 'Artificial intelligence and climate change' (2020) 37 JREG 890; PK Lin, 'The cost of training a machine: lighting the way for a climate-aware policy framework that addresses Artificial Intelligence's Carbon Footprint Problem' (2022) 34 Fordham Envtl L Rev (forthcoming), Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4066935 (accessed 12 January 2023).

8 See eg C Heinze, 'Patent law and climate change – do we need an EU patent law directive on clean technology?' (2021) 70 GRUR Int 554.

9 T Pihlajarinne 'Repairing and Re-Using from an Exclusive Rights Perspective – Towards Sustainable Lifespan as Part of a New Normal?' in AO Rognstad and B Ørstavik (eds) *IP and Sustainable Markets* (Elgar, 2021) 81; A Tischner and K Stasiuk, 'Spare parts, repairs, trade marks and consumer understanding' (2023) 54 IIC *26*.

10 MP Schinkel and L Treuren 'Green Antitrust: (More) Friendly Fire in the Fight against Climate Change' in S Holmes, D Middelschulte, and M Snoep (eds) *Competition Law, Climate Change & Environmental Sustainability, Concurrences* (2021). Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3749147 (accessed 22 March 2023).

11 See eg C Marquis, MW Toffel and Y Zhou, 'Scrutiny, norms, and selective disclosure: a global study of greenwashing' (2016) 27 Organ Sci 483; K Park 'Green Trademarks and the Risk of Greenwashing' WIPO Magazine (22 December 2022). Available at https://www.wipo.int/wipo_magazine/en/2022/04/article_0006.html (accessed 22 February 2023).

12 See eg Daria Kim, 'No one's ownership as the status quo and a possible way forward: A note on the public consultation on Building a European Data Economy' (2018) 13 J Intell Prop L & Pract 154; and J Drexl et al, 'Data Ownership and Access to Data - Position Statement of the Max Planck Institute for Innovation and Competition of 16 August 2016 on the Current European Debate' Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2833165 (accessed 22 February 2023); German Ministry of Justice and Consumer Protection and Max Planck Institute for Innovation and Competition (eds), Data Access, Consumer Interests and Public Welfare (Nomos 2021).

13 One explanation about data being locked in automotive sector is offered by Daniel Gill 'The Data Act Proposal and the Problem of Access to In-Vehicle Data and Resources' Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4115443 (accessed 12 January 2023).

in vast databases of device manufacturers who deem such data as their proprietary asset and have no incentive to share it with the owners of sensorized devices nor with third parties who might be interested in creating new value (applications and services) with such data.[14]

In this article, we proceed from the observation that a lot of 'green data' are captured by a wide variety of sensors and that the sensorization of everything is one of the defining features of digital transformation. Such sensors are embedded in industrial machinery, smart internet of things (IoT) devices used by consumers and businesses and personal sensors that individuals interact with in every aspect of their daily lives (eg wearable fitness devices or IoT devices at home). The trend of sensorization and the quantity and quality of data generated will continue to increase exponentially. Finding a bit way to leverage the value of such data is a defining challenge of our age.

We argue that one of the key aspects of building a more sustainable future and combating climate change is to unlock the green value of these data. In Section 2, we provide some background context to our argument, namely, the sensorization of everything, the concept of unlocking the green value of data and a user-held data model, which we suggest represents a better way of leveraging such unlocked value. In Section 3, we explore the potential of the user-held data model and possible applications in the context of smart cities and agriculture, empower consumers to make 'greener' choices in their daily lives and facilitate the 'green data economy'. In Section 4, we conclude by identifying some regulatory principles for unlocking green data on a global scale and activating such data to meet carbon equation objectives. The main argument is that more green data need to be unlocked and activated to reach the carbon equation expeditiously.

2. Unlocking the value of green data

Over the past decade, cloud computing, Big Data and data analytics technologies have been at the heart of innovations that permeate all sectors of the economy. Data have become a defining feature of the contemporary world, ie sensor-equipped devices and artificial intelligence (AI) applications built on top of such sensor-generated data open new opportunities that could also help alleviate the effects of climate change. The trend of sensorization, collection of data and development of AI solutions happens in all sectors of the economy (mobility, real estate, construction, agriculture, health, energy, etc).

In this section, we argue that the key aspect of building a more sustainable future is to unlock the green value of sensor-generated data. Here, we introduce the so-called 'user-held data' model that is based on a human-centric approach to data and the idea that user-generated data should belong to users.[15] We focus on explaining the technical architecture behind the user-held data model and the innovation breakthrough that will happen in an open data ecosystem where everyone can collect their data in their own personal data cloud and where each individual's data are private by default.

2.1. The sensorization of everything

Sensors are coming into every aspect of our daily lives. This trend was proven at the Consumer Electronics Show 2023 (CES 2023)—the largest global consumer electronics show of the year which took place in Las Vegas in January 2023. More than 3200 companies presenting their latest innovations were in the category of 'IoT' and 'AI.'[16] In this regard, CES 2023 was a 'reality check' confirming that personal sensors, user-generated data and software built on top of such data (eg AI tools and AI assistants) are becoming ubiquitous.[17]

The current sensorized data ecosystem can be categorized into three segments or layers.[18] First, sensors are being integrated into all possible daily scenarios to measure the performance of devices and engines as well as monitor various human activities. Sensors are vital in measuring the temperature of buildings and engines, the operation of robots and digital twins in fulfilling various tasks and making necessary adjustments. For instance, a consumer vehicle is loaded with sensors and electronics that measure more than 200 data points (eg when a car is unlocked, the engine starts, speed and brakes, surrounding objects, etc).

Second, many consumer and business devices have a sensor embedded in the device and are accompanied by

14 Ibid. For an overview of data silos in the IoT sector, see this excellent paper by Wolfgang Kerber, 'Governance of IoT data: Why the EU Data Act Will not Fulfil Its Objectives' (2022) 72 GRUR Int 120.

15 P Jurcys, C Donewald, M Fenwick, M Lampinen, V Nekrošius, and A Smaliukas, 'Ownership of User-Held Data: Why Property Law is the Right Approach' (2021) 34 Harv J L & Tech Digest, Available at https://jolt.law.harvard.edu/digest/ownership-of-user-held-data-why-property-law-is-the-right-approach (accessed 12 January 2023).

16 Wired 'CES 2023 in Photos: Welcome to Gadgetville, USA' Wired (8 January 2023). Available at https://www.wired.com/gallery/ces-2023-photos-roger-kisby/ (accessed 12 January 2023) and P Jurcys 'Predicting the Future: Reflections from CES 2023' Available at https://www.prifina.com/blog/230118 (accessed 12 January 2023).

17 RL Kaliouby 'Wellness on Wheels Is Possible' Wired (8 January 2023). Available at https://www.wired.com/story/wellness-cars-driving-transportation/ (accessed 12 January 2023).

18 See Jurcys et al (n 15).

an application that usually provides a dashboard of real-time data and data-driven trends. Usually, such data are 'co-generated' by the user of the device and sent to the centralized server of the app maker, where the data are processed and the generalized results are sent back to the dashboard of the user.

Finally, data are utilized in various scenarios by companies that are training machine learning algorithms to solve specific problems. For instance, in business-to-business scenarios, there are companies that are developing digital twins to assist businesses in simulating technical trials, facilitating research and development or conducting clinical trials of medical drugs.[19] There are also companies that build various AI tools to train language, speech or vision tools that can be applied in business scenarios and daily consumer transactions.

Assuming that the trend of massive adoption of sensors of the amount of sensor-generated data is going to increase, it is important to find a solution for how such sensor-generated data could be utilized in an efficient and meaningful way for the benefit of the greatest number of stakeholders on the market.

Currently, the digital infrastructure is *enterprise* and *product-centric*: consumers and business users must have an account for every product they use.[20] In practical terms, such a product-centric data architecture means that device makers and service providers collect data about each user. The current centralized data architecture is not only inconvenient for customers ('application fatigue') but also raises numerous technological (interoperability) challenges.[21] More importantly, a product-centric approach to data means that much of the user-generated data are locked in walled gardens of sensor makers and device providers.[22]

The agricultural sector is an illustrative example of data such being locked in. Although most data originate in farms, such agricultural data are harnessed and processed by third-party makers of sensors, machines and service providers.[23] Data are predominantly collected and stored in closed, siloed databases of manufacturers of sensorized devices and machines used in farming and agriculture; farmers also rely on the services of third-party service providers.[24] Such a data ecosystem results in information and power asymmetries where farmers have limited control over the data they themselves are generating through their agricultural activities. In legal practice, farmers have limited power to adjust draft contracts with service providers, especially when it comes to using and sharing of farmer-generated data with third parties with whom the farmer has no relationship. When it comes to maintenance and repairs of the equipment used in agriculture, there is an increased push in Europe and elsewhere to introduce the 'right to repair'.[25] As of today, in many cases, repair and maintenance of sensorized products are limited to 'authorized' repair shops. This leads to increased costs and limited competition in the repair aftermarket. These circumstances in the agricultural market explain why the European Union (EU) Commission is pushing for more agricultural data sharing.[26]

2.2. Recent regulatory initiatives to unlock green data

Since the adoption of the GDPR[27] in 2018, the EU has been exploring new ways to promote access and flow of data that are beneficial not only to the largest Silicon Valley tech giants and large manufacturers of sensorized products but also—and crucially—to individual users and businesses. The EU estimates that the size of

19 See eg, https://www.3ds.com/products/medidata (accessed 12 January 2023).

20 See 'Product-Centric vs Customer-Centric: What's the Difference, and How to Switch to a Customer-Centric Approach?' Userpilot (7 June 2022). Available at https://userpilot.com/blog/product-centric-vs-customer-centric/ (accessed 12 January 2023).

21 For an overview of how various stakeholders perceive the current data ecosystem and what are they key pain-points, see P Jurcys and M Lampinen 'Principles of Data Privacy in California: Study of Industry Reactions and Comments to the Proposed CCPA Regulations and User-Centric Perspectives' Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3601948 (accessed 12 January 2023).

22 One recent study has shown that, in 2021, there were 178 cases of data marketplaces for enterprises, among which two were dedicated to agricultural data, four covered data regarding the connected car and the automotive industry, two on sensor data and nine on other B2B data. See Simon et al, Definition and analysis of the EU and worldwide data market trends and industrial needs for growth, 2021, p. 26; the database is accessible under https://doi.org/10.4121/14679564.v1.

23 H Schweitzer et al., 'Data Access and Sharing in Germany and in the EU: Towards a Coherent Legal Framework for the Emerging Data Economy' p. 155. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4270272 (accessed 12 January 2023).

24 Ibid.

25 See eg European Commission 'Sustainable Consumption of Goods – Promoting Repair and Reuse' Available at https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13150-Sustainable-consumption-of-goods-promoting-repair-and-reuse_en (accessed 12 January 2023); LC Grinvald and O Tur-Sinai, 'Intellectual property law and the right to repair' (2019) 88 Fordham L Rev 63.

26 It is believed that the EU's 'Code of Conduct on Agricultural Data Sharing by Contractual Agreement' could help tackle several of these issues eg by clarifying roles and who has control of the data or in providing a framework for data portability and EU Agricultural data space: Available at https://digital-strategy.ec.europa.eu/en/events/information-session-common-european-agricultural-data-space (accessed 12 January 2023).

27 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119/1.

the data economy in 2027 will be approximately 829 billion euros.[28] In light of the fact that increasing amounts of data are generated every day, the EU has been developing additional measures that might level the playing field in the data market.

In February 2020, the EU Commission published a European strategy for data, where it laid out its ambition to make the EU the leading role in the global data economy.[29] To achieve this goal, the EU is now adopting forward-looking laws and regulations for various 'data spaces' (eg health, agriculture, mobility, etc) and establishing a framework for responsible AI.[30] The EU primarily focuses on opening the data market: making data accessible to various stakeholders who can build new types of products and services on top of data. Hence, data accessibility, portability and interoperability have become the foundational principles on which the EU aims to build its leadership in the global data market.

For green data purposes, one of the key EU instruments is the proposed European Data Act,[31] which seeks to strengthen the EU's position on the global technology market by opening up data to boost innovation.[32] The proposed Data Act, which will apply to the manufacturers of IoT devices, connected products and services, aims to lay more detailed rules to make high-quality data available and flowing in all sectors of the economy, including the data 'aftermarket.' The EU Data Strategy seeks to foster economic growth, competitiveness and innovation, maintain high data protection standards, maintain the balance between human rights and take into consideration the impact of the economy on climate.

Considering the paramount importance of data in dealing with carbon emissions, it is possible to see how the proposed EU Data Act fits within a broader set of policy objectives of the EU to meet its environmental objectives. More specifically, the European Green Deal, adopted in 2019, aims to make Europe the first carbon neutral continent in the world.[33] The current strategy of the EU is to invest heavily into environmentally friendly technologies and sustainable growth policies and reach carbon neutrality by 2050 at the latest. The carbon reduction goals align with other economic and social aspirations: to make energy prices more affordable and transportation and hospitality smarter and more energy efficient, create new jobs and improve the overall quality of life.[34] The EU Green Deal thus accompanied by the Investment Plan totalling over 1 trillion euros into this 'green' transition, which also incorporates digital and data infrastructure.[35]

One of the more important practical objectives of the proposed Data Act is to unlock the data currently collected by and harnessed by the manufacturers of IoT devices. This covers smart IoT devices in manufacturing and industrial applications (such as robotics) and consumer-facing IoT devices (eg smart watches, fitness and sleep trackers). Article 35 of the proposed Data Act is of crucial importance: it stipulates that IoT device makers must give users access to the databases that contain data obtained from or generated using the IoT device or a related service.[36] More importantly, Article 35 clarifies that manufacturers of smart IoT devices do not have exclusive rights over the databases containing data obtained from or generated by using a product or related service.

The Data Act makes a bold step forwards and strips away database makers of such exclusive rights over user-generated data and clearly states that consumers and businesses have the right to access data generated by the products or related services they own, rent or lease (Article 3). Furthermore, manufacturers and designers of IoT devices and accompanying services will have to design the products to make the data easily accessible by default (Article 3(1)). Besides, manufacturers of IoT devices will have to be transparent on what data will be accessible and how

28 European Data Strategy, see: Available at https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en (accessed 12 January 2023).

29 Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A European strategy for data, COM/2020/66 final (18 February 2022).

30 See Proposal for a regulation—The European Health Data Space, COM(2022) 197/2; and Proposal for a Regulation of the European Parliament and of the Council Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, SEC(2021) 167 final.

31 Available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0068 (accessed 12 January 2023).

32 See Available at https://digital-strategy.ec.europa.eu/en/library/data-act-proposal-regulation-harmonised-rules-fair-access-and-use-data (accessed 12 January 2023).

33 'A European Green Deal' Available at https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en#documents (accessed 12 January 2023).

34 Norton Rose Fulbright 'The EU Green Deal explained' Available at https://www.nortonrosefulbright.com/en/knowledge/publications/c50c4cd9/the-eu-green-deal-explained (accessed 16 February 2023).

35 See Available at https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_24 (accessed 16 February 2023).

36 The scope of access is limited to the 'observed' data. For a more detailed analysis, see A Metzger and H Schweitzer, 'Shaping Markets: A Critical Evaluation of the Draft Data Act' (2023) ZEuP 1/2023 42 and J Drexl et al, 'Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act)' Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4136484 (accessed 12 January 2023).

to access them. The Data Act also will create an obligation for the data holder to make such data available to third parties upon the user's request. Users will be entitled to authorize the data holder to give access to the data to third-party service providers, such as providers of aftermarket services.

2.3. Key principles of a user-held data model

We are at the turning point where the human-centric technology ecosystem becomes real and where individuals are actually in control of their own data.[37] The user-held data model refers to a new data infrastructure which is built around an individual consumer who can collect data from various personal data sources real time and have actual ownership and control of such data. This human-centric approach to personal data is based on the assumption that user-held data should be private by default and not shared with anyone unless the user decides otherwise. Rather, third-party applications should come to the user and run on top of user-held data. The user-held data model is grounded the foundational principles of GDPR (consent, transparency, security of personal data, data minimization, purpose limitation and data portability).[38]

The idea of individuals owning their data is not new. In the early days of the internet, entrepreneurs working at up-and-coming companies thought that users' data should belong to them. Craig Federighi, one of the executives at Apple, once explained that 'when Apple was founded, the proposition was that this is your personal computer; this is your own data. As the world has evolved, we have continued to think that the data you create, the things you do with your computer, those are yours and you should be able to control it.'[39] Reputation.com, for example, was built on the same idea that individuals should be able to own and control their data.[40] However, as internet search and traffic started to concentrate around several central nodes in the web, this notion of user-owned data was marginalized.

Federated cloud architecture and edge computing technologies have now matured to empower individuals with their data.[41] There are companies that are building solutions based on the so-called 'user-centric, user-held data' model which contains two defining features.[42] First, each individual user has their personal data cloud instance where they can collect data from personal sensors, wearable devices and online data sources (eg location from Google Maps, shopping history from Amazon, activity while using social media services, etc). Each personal data cloud has an embedded software robot that unifies data from different data sources. It is important to note that each user's data are private by default and nobody other than the individual has access to the personal data cloud (unless the individual gives access to it).[43] Second, developers and companies can build new applications that 'come to the user' and run locally (ie 'on top of' user-held data) instead of sending data to centralized siloed environments of Google, Apple or any other company providing a sensorized device.

A user-held data model opens new opportunities with personal data: developers can easily build applications without having to solve complex data (back-end) problems. Federated data architecture where users own and control their personal data frees developers from the hefty burden of complying with data privacy regulations because apps run locally on top of user-held data and generate value on the user side (instead of being centralized in a service provider's platform). Such user-centric technology infrastructures built with a private-by-default approach to personal data reveal promising signs of actual implementation of fundamental human rights and provide a better basis for drafting new regulations for data and AI.

From a legal point of view, the user-centric, user-held data architecture emanates and is in line with the core privacy-by-design and data privacy principles enshrined in GDPR, California Consumer Privacy Act[44] and other

37 Jurcys et al (n 15); P Jurcys, M Corrales Compagnucci and M Fenwick, 'The future of international data transfers: managing legal risk with a 'user-held' data model' (2022) 46 CLSR 1.

38 For a more detailed analysis, see P Jurcys 'User-Centric, User-Held Data Model: Key Principles' Available at https://medium.com/prifina/user-centric-data-model-key-principles-d02a69cf45d0 (accessed 12 February 2023).

39 Michael Grothaus 'Exclusive: Apple's Craig Federighi on WWDC's big privacy upgrades and beyond' Fast Company (22 June 2020). Available at https://www.fastcompany.com/90518458/exclusive-apples-craig-federighi-on-wwdcs-big-privacy-upgrades-and-beyond (accessed 22 February 2023).

40 M Fertik and D Thompson, *The Reputation Economy* (Crown Business, New York 2015).

41 M Fenwick and P Jurcys, 'From cyborgs to quantified selves: augmenting privacy rights with user-centric technology and design' (2022) 13 JIPITEC 20.

42 See eg Available at www.prifina.com (building a human-centric data ecosystem where each user's data are private by default) or Tim Berners-Lee's Inrupt: Available at https://www.inrupt.com/ (building a consumer data portability ecosystem for enterprises where individual consumers are objects in the ecosystem).

43 P Jurcys, C Donewald, J Globocnik, and M Lampinen, 'My Data, My Terms: a proposal for personal data use licences' (2020) 33 Harv JL & Tech Digest 9–12.

44 The California Consumer Privacy Act of 2018 (CCPA). For more information, see Available at https://oag.ca.gov/privacy/ccpa (accessed 12 January 2023).

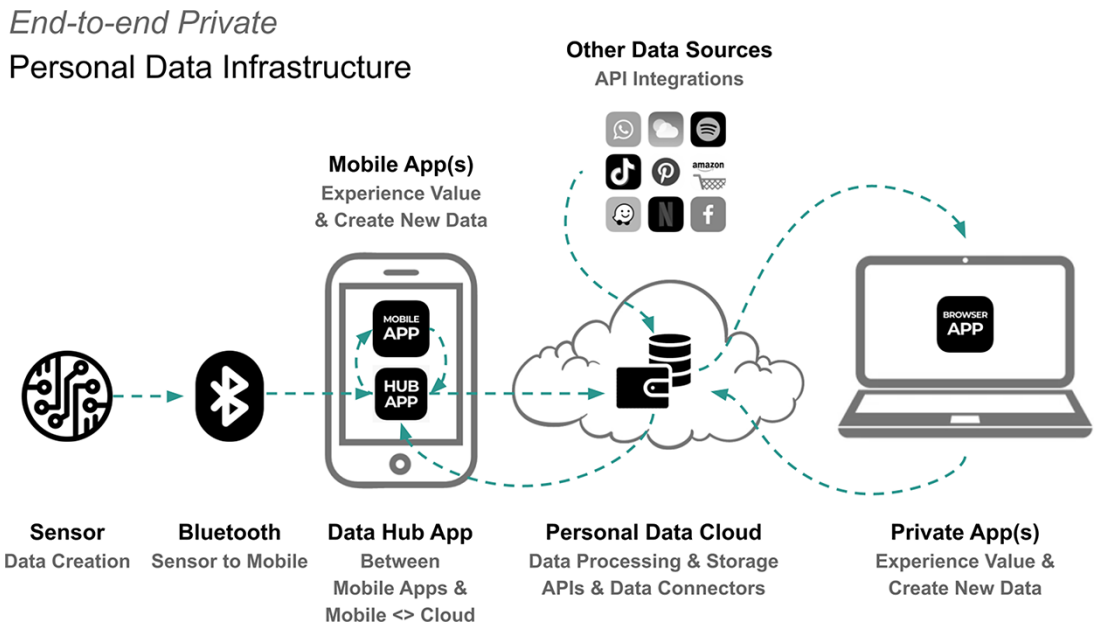

Figure 1. The personal data infrastructure of the future.

recent data privacy regulations.[45] Apps that run on top of user-held data are operating under the terms set forth by the user.[46] This user-centric approach to data constitutes a remarkable improvement for user experience: rather than opting out from various types of data collection practices superimposed by service providers, individuals actually are able to effectively determine whether they are willing to grant access to certain datasets to third-party applications which they download to their personal data clouds ('opt-in' environment). The user-held data model also helps application developers comply with the principle of data minimization (Article 5(1)(c) GDPR) because, by default, user-held data stay locally with the user and applications run 'on top of' such user-held data—the service provider (application developer) does not have to hold any data on its servers (See Fig. 1).

3. Unlocking green data as a resource for more sustainable living

The collection and use of green data from various sensors and online data sources is one of the elements in designing tools and solutions for tackling climate change.[47] In this section, we explore three possible dimensions or contexts where user-generated sensor data combined with the user-held data model outlined earlier can help 'activate' green data and, by doing so, trigger or nudge the creation of innovative climate technologies. In particular, we focus on the utility of a user-held data model to help people adopt green lifestyles,[48] to facilitate sustainability in green cities[49] and to build green innovation ecosystems[50]

45 A Cavoukian 'Privacy by Design: The 7 Foundational Principles' Available at https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf (accessed 22 February 2023).

46 See Fenwick and Jurcys (n 41).

47 See eg G Noto La Diega and E Derclaye, 'Opening Up Big Data for Sustainability: What Role for Database Rights in the Fourth Industrial Revolution?' in OA Rognstad, T Pihlajarinne and J Mähönen (eds) *Promoting Sustainable Innovation and the Circular Economy: Legal and Economic Aspects* (Routledge, Abingdon 2022). Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4220534 (accessed 12 January 2023); B Crona and E Sunsdström, 'Sweet spots or dark corners? An environmental sustainability examination of big data and AI in ESG' in T Rana, A Lowe and J Svanberg (eds) *Handbook of big data and analytics in accounting and auditing* (Springer, Singapore 2023).

48 To illustrate our point, we primarily refer to the literature on behavioural economics and environmental nudges: M Binder and AK Blakenberg, 'Green lifestyles and subjective well-being: More about self-image than actual behavior' (2017) 137 J Econ Behav Org 304; LA Reisch et al, 'Mitigating climate change via food consumption and food waste: a systematic map of behavioral interventions' (2021) 279 J Clean Prod 123 at 717; Alessandro Bucciol, Alessandro Tavoni and Marcella Veronesi (eds), *Behavioural Economics and the Environment: A Research Companion* (Routledge, Abingdon 2023).

49 See eg ME Kahn 'Sustainable and Smart Cities' Policy Research Working Paper No. 6878. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2439699 (accessed 22 February 2023); BM Frischmann, MJ Madison and MR Sanfilippo (eds), *Governing Smart Cities as Knowledge Commons* (CUP, Cambridge 2023).

50 See eg P Engelke, M Jackson and R Bell 'Mapping Green Innovation Ecosystems: Evaluating the Success Factors for the World's Leading Greentech-Innovation Centers' Available at https://www.atlanticcouncil.org/wp-content/uploads/2021/01/Greentech-Issue-Brief-2021-1.pdf (accessed 22 February 2023).

3.1. Green lifestyles

Recent socio-economic studies show that individual consumers value a clean environment (water and air) and are increasingly willing to utilize the data they generate by using various personal sensors and the data they generate in the digital environment.[51] How could such user-generated data empower individuals to make more conscious decisions in their daily lives to raise awareness about more sustainable lifestyles? The user-held data model combined with the insights from behavioural scientists about nudges and disclosures could help spur innovation in this domain.[52]

We are all very accustomed to seeing food labels that disclose various information about the products we are buying, eg how much sugar, fat, carbs and calories they contain. And, of course, it is ultimately up to every consumer to decide if they want to read the label. However, making the information available in a clear and transparent form is a big step forward in helping people make informed choices about the food they purchase.[53] Apple's 'privacy labels' are similar disclosures about how mobile phone applications utilize our data.[54] Many of us may not check this information before downloading apps, but it helps app users check if they are happy with how their data are accessed, collected and used.

One of the major benefits of the user-held data model is that it allows the building of new types of applications that can correlate data from various data sources of an individual user. Take mobility as an example. Nowadays, there are examples where air transportation companies are exploring ways to inform consumers how their use of certain services contributes to the global carbon equation. For instance, some airlines are giving consumers the option to purchase more expensive tickets and subsidize the use of less polluting diesel during the flight. If we think about personal IoT data, what kinds of applications could be created that would provide insights based on personal sensors? The possibilities seem endless.

A developer can build a 'My Carbon Footprint' application that correlates data from user's location data from Google Maps and Apple Maps, how often the user uses public transport (bus, train and taxi), how often a user is cycling and driving his own car and how often they travel by plane. A 'My Carbon Footprint' app of this kind shows correlations between different data points and helps people better understand their daily/monthly carbon footprint. Such an app can also include additional personalized insights that make end-users more aware and responsible for how their actions affect the environment.

Similar applications can be built on user-held data about using energy at home (eg how much energy is consumed in total, what percentage of energy is coming from renewable sources and offer recommendations on whether certain heating/cooling IoT devices could be more energy efficient). Additional insights can be provided by app developers and data scientists who correlate data from different categories of data sources (eg how outdoor exercise activities decrease/increase energy consumption at home).

In the domain of personal health and wellness, there are many valuable applications that can be built on top of user-generated data from fitness devices to help people choose healthier lifestyles (eg an app that could combine personal fitness goals with public weather data showing which jogging route in a city is less polluted, etc).

Finally, the technological architecture of user-held data could be combined with rich findings of behavioural scientists to offer better insights in the form of disclosures, feedback mechanisms and nudges to help people rely on their green data to navigate daily situations to lead more sustainable lifestyles.[55]

3.2. Green cities

Current smart city solutions are based on one premise: the need to collect lots of data to develop and improve 'smart' city services. Such data are collected in various public settings (eg streets, highways, parking lots and public squares) utilizing various sensors (eg parking metres) and surveillance tools (eg video cameras and facial recognition technologies). To collect such data, municipal authorities engage with many private vendors who are entrusted with complex tasks to manage networks, infrastructure and big data. Personal data are becoming a privacy liability for both public and private organizations. In some cases, certain organizations collecting personal data in centralized servers make them uninsurable. This is particularly true of municipal and public organizations; they have much lower risk

51 AG Winegar and CR Sunstein, 'How Much Is Data Privacy Worth? A Preliminary Investigation' (2019) 42 J Consum Policy 425; S Dinan 'Shoppers' vast giveaways of data help cheapen personal privacy' Washington Times (25 December 2022). Available at https://m.washingtontimes.com/news/2022/dec/25/its-christmas-season-how-much-your-data-did-you-gi/ (accessed 12 January 2023).

52 About the practical aspects of voluntary and mandatory disclosures, see CR Sunstein, *The Cost-Benefit Revolution* (MIT 2018) ch 7 and 9.

53 CR Sunstein, 'On mandatory labeling, with special reference to genetically modified foods' (2017) 156 Univ PA Law Rev 1043.

54 Apple's privacy labels for apps are available here: Available at https://www.apple.com/privacy/labels/ (accessed 22 February 2023) for an evolution of privacy labels see P Jurcys, 'Privacy Icons and Legal Design' Available at https://towardsdatascience.com/privacy-icons-4ca999a6f2db (accessed 22 February 2023).

55 RH Thaler and CR Sunstein *Nudge: The Final Edition* (Penguin, New York 2021) 281–286, 301–304; M Fenwick and P Jurcys (n 41) 25–26.

tolerance. Due to the technical and compliance burden, many public and private entities are simply unable to leverage significant personal data centrally. Oftentimes, they have to make a difficult choice: limit data collection and use or find another way of using it, such as the user-held data model.

The arrival of 5G networks and edge computing technologies is seen as key driving force to facilitate smart city projects, update infrastructure and build new solutions based on huge amounts of the data collected. Data are at the heart of many smart city interactions: optimizing traffic and public transport and distributing energy consumption, power and water supply, waste management, crime control, health care and community services. All of these have a direct impact on the environment: improving air quality, waste of food and energy, noise pollution and designing responses to climate disasters.

Although smart city projects aim to optimize the use of public resources, such projects also raise surveillance concerns and impinge upon the citizens' privacy rights. The user-held data model could be especially useful to move away from 'Big Brother'/'Uncle Sam' scenarios and develop more citizen-focused or human-centric solutions.[56] It has been suggested that citizen-centric smart city solutions should be built on user-generated and user-held data where data activated on the user side are more suitable for better services, optimize their mobility, help save time and money and improve the quality of life in a city.

The user-held data model might open the gates for cities and service providers to build various applications that combine private (user-generated) and public data. For example, user-generated mobility data from smart wearables could be combined with public data (eg weather and air quality data) to offer individuals healthier walking routes. Similarly, public and user-generated energy consumption data could be utilized to better manage energy distribution during peak hours.

In a post-Covid environment, many applications could be created to increase the utility of idle resources (eg unused office spaces and empty parking lots) or use privacy-preserving data access frameworks to build more resilient cities (eg manage access to buildings and office spaces without exposing sensitive employee data). Taken as a whole, the combined use of private and public green data could generate more value for society.

56 Jouko Ahvenainen 'Smart City Models Often Overlook One Key Component – the People in Them' Available at https://disruptive.asia/smart-city-models-overlook-people-in-them (accessed 12 January 2023); UN Habitat 'People-Centered Smart Cities' Available at https://unhabitat.org/programme/people-centered-smart-cities (accessed 12 January 2023).

3.3. Green innovation ecosystems

From a more macro/ecosystem perspective, unlocking green data can provide an important catalyst for entrepreneurship and sustainable innovation within a particular region or territory. New green innovation ecosystems can emerge if suitable institutional, technological and regulatory mechanisms are in place.

The aforementioned EU Data Strategy and related data and AI regulations combined with the new federated data and machine learning solutions are seen as major catalysts for innovation with green data. Policymakers on a national and regional level need to cooperate closely with other stakeholders (universities, tech giants, public interest groups, venture capital industry, as well as emerging technology companies) to explore mutually beneficial objectives and adopt measures to further sustainable innovation that aims to activate as much green data as possible.

Such activities can contribute to the ripple effect where successful test cases of using green data will be gradually replicated in different industries and help advance towards reaching the carbon equation. Continuous interaction among various players in the green innovation ecosystem should facilitate mutual trust and lead to a positive feedback loop around green data and individual lifestyles, cities and ecosystems.

4. Paths forward: principles for unlocking the green value of data

In thinking about regulatory interventions and building the green data ecosystems of the future, the following three principles can provide guidance for policymakers: (i) emphasis on personal autonomy and empowerment, (ii) promoting interoperability and (iii) regulatory experimentation.[57] The emerging focus on data access and user-held data solutions can have the biggest societal impact if the need to protect individual autonomy is considered as prevailing over the monopolistic interests of corporate stakeholders in the IoT market. Furthermore, a human-centric approach to data leads to the creation of more ethical products and services and accelerates the emergence of standards that are needed for interoperability purposes. Finally, we argue for more experimentation in the regulatory domain—an approach that is likely to help more innovative green data products enter the market.

57 For a general overview see Adam Thierer, 'The internet of things and wearable technology: addressing privacy and security concerns without derailing innovation' (2016) 21 Rich J L & Tech 6.

4.1. Emphasizing personal autonomy and empowerment

The advancement of cloud computing and data processing technologies marks the transition from the centralized surveillance information environment towards a more decentralized internet, where individuals have more control and ownership over their data. In practical terms, this shift means that individuals will regain their digital autonomy and be able to make more conscious choices over how their data are used and who has access to it. As a comparison, in the current enterprise and product-centric environment, individuals can only request more transparency from service providers over how they use customer data, and users can have only one option—to opt out of certain companies' data collection practices. Unfortunately, today, data privacy does not mean that users' data are private.

From a human-centric perspective, one may question why the users should bear the hefty burden of figuring out how companies use their customer data. Is it possible to turn the tables and create an environment where, by default, users are in control over their own data and where no one has access to user data without prior explicit user consent?[58] This is exactly what the user-held data model is about: collecting data into one's personal data cloud is just one first step. Having apps come to the user and run locally (on top of user-held data) means that at least the largest chunk of sensitive personal data remains private and the value from that data also remains on the user's side.

Giving data control back to individuals is a precondition to nudge individuals to 'unlock' data for green projects. Such green value from user-generated data could take shape even in the form of disclosures and labelling: being able to understand how their behaviour and lifestyle affect the environment and contribute to climate change could be transformative, especially if we think about the climate change impact on a global scale.[59]

4.2. Promoting interoperability

With sensors becoming ubiquitous, interoperability remains one of the underlying technical challenges. As we discussed earlier, the network of physical devices that are connected to the internet and can collect, send and receive data is growing. Interoperability between these devices is important; otherwise, the full potential of IoT will not be realized. The current enterprise and product-centric data architecture provides little incentives for device manufacturers to cooperate in developing common standards that promote interoperability.

Interoperability can be facilitated by using a human-centric approach to data and open-source standards. A human-centric approach to data facilitates interoperability by nudging data scientists and developers to design IoT systems by considering how different devices can work together to meet the user's needs. This means that the data generated by various IoT devices should 'speak to one another' with the goal of benefiting the end-user. For example, by understanding the user's daily routine, different devices can be programmed to work together to automatically perform certain tasks, such as turning on the lights and adjusting the thermostat when the user arrives home. This human-centric approach to data and IoT systems should be embraced by various stakeholders, including the Institute of Electrical and Electronics Engineers and the Internet Engineering Task Force, which play a role in developing and promoting standards for IoT device interoperability.

4.3. Regulatory experimentation

Over the last few decades, the theory and practice of regulation have focused on improving the ability of regulators to respond to changing industry practices and on improving relationships between regulators and regulated entities.[60] So-called dynamic or responsive regulation responds to changing industry practices through feedback effects and enhanced information for developing regulatory forms.[61]

What is particularly important is that within the framework of these more dynamic and agile legal frameworks, regulatory decisions should not be thought of as 'final events' that are made once and for all and from which we then all move on. Rather, we should think of regulatory choice as a form of what French sociologist, Michel Callon has called 'measured decision-making,' ie regulation involves open-ended and highly contingent

58 See eg MA Lemley 'The Benefit of the Bargain' Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4184946 (accessed 12 January 2023); DJ Solove 'Murky Consent: An Approach to the Fictions of Consent in Privacy Law' Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4333743 (accessed 12 January 2023) and WEF 'Data for Common Purpose: Leveraging Consent to Build Trust' Available at https://www.weforum.org/whitepapers/data-for-common-purpose-leveraging-consent-to-build-trust/ (accessed 22 February 2023).

59 Sunstein (n 53).

60 See J Black, M Hopper and C Bland, 'Making a success of principles-based regulation' (2007) 1 Law Financ Mark Rev 191.

61 M Fenwick, WA Kaal and EPM Vermeulen, 'Regulation tomorrow: what happens when technology is faster than the law' (2017) 6 Am Univ Bus Law Rev 561; M Fenwick, EPM Vermeulen, 'The Future of Finance: Why Regulation Matters' in KT Liaw (ed) *The Routledge Handbook of Fintech* (Routledge, Abingdon 2020).

selections that are merely one stage or moment in a longer process and not the final word on a particular issue.[62]

As such, to facilitate the testing and adoption of new technology frameworks for green data, regulators must embrace contingency, flexibility and openness to learning. The justification for this new openness derives from the contingency of the fast-moving business environment in which regulators now operate. Any regulatory choice ultimately remains open ended, leaving space to incorporate new knowledge, discoveries and claims. The need for finality, Callon suggests, is overstated and is more typically the product of expediency and habit than genuine necessity.

This new emphasis on contingency and iteration in regulation is a function of the fast-changing environment in which regulation now operates, specifically globalization and technological change. This complex new operating environment creates a challenge for policymakers, and one response has been a willingness to innovate in regulatory schemes and remove the temporal horizon that has traditionally defined decision-making. Under such circumstances, policy experimentation, ie testing different regulatory schemes and then comparing the results, becomes attractive and seems important in the context of environmental policy.[63]

5. Concluding remarks

In this article, we analysed how user-generated data can be utilized to address the looming climate change problem. We emphasized that the role of data has received only marginal attention from various stakeholders who seek to adhere to multiple ESG principles and objectives. We introduced the human-centric approach to data which shifts the attention from the enterprise-first use of large datasets towards local data generated by everyone. Furthermore, we introduced the notion of green data and green value that could be created by utilizing user-generated data. We argued that such user-generated data and user-held data models could unlock the value of data for individuals who could adjust their lifestyles to alleviate the impact of their behaviour on the environment.

Finally, we introduced three normative principles that regulators and policymakers might usefully embrace in pursuing these objectives. The value of green data might be unlocked in regulators' focus on the needs of individuals. A user-held data model should receive more attention because it resonates with the interests and incentives of all stakeholders in the green data economy of the future. We also encouraged a softer approach to law-making and regulatory interventions via the proposed increased use of regulatory experimentation. Such more contingent regulatory measures can facilitate faster market adoption of technologies that utilize green data. More generally, we invite regulators to adopt a more open and compassionate mindset towards innovators who seek to find ways to reduce carbon emissions: experimentation and a trial-and-error method can also facilitate innovation in climate technologies.

62 M Callon, P Lasoumes and Y Barthe, *Acting in an Uncertain World: An Essay on Technical Democracy* (MIT Press 2009).

63 See Ministry of Economy, Trade and Industry of Japan, 'Governance Innovation Ver. 2: A Guide to Designing and Implementing Agile Governance' Available at https://www.meti.go.jp/press/2021/07/20210730005/20210730005-2.pdf (accessed 12 January 2023).