Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Environmental Sustainability

The role of data in transformations to sustainability: a critical research agenda

João Porto de Albuquerque[1], Liana Anderson[2], Nerea Calvillo[3], Jon Coaffee[4], Maria Alexandra Cunha[5], Livia Castro Degrossi[5], Giovanni Dolif[2], Flavio Horita[6], Carolin Klonner[7], Fernanda Lima-Silva[5], Victor Marchezini[2], Mario Henrique da Mata Martins[5], Diego Pajarito-Grajales[1], Vangelis Pitidis[1], Conrado Rudorff[2], Nathaniel Tkacz[3], Rachel Traijber[2] and Alexander Zipf[7]

This article investigates the role of digital technologies and data innovations, such as big data and citizen-generated data, to enable transformations to sustainability. We reviewed recent literature in this area and identified that the most prevailing assumption of work is related to the capacity of data to inform decision-making and support transformations. However, there is a lack of critical investigation on the concrete pathways for this to happen. We present a framework that identifies scales and potential pathways on how data generation, circulation and usage can enable transformations to sustainability. This framework expands the perspective on the role and functions of data, and it is used to outline a critical research agenda for future work that fully considers the socio-cultural contexts and practices through which data may effectively support transformative pathways to sustainable development.

Addresses
[1] Institute for Global Sustainable Development, University of Warwick, UK
[2] National Center for Monitoring and Early Warning of Natural Disasters—CEMADEN, Brazil
[3] Centre for Interdisciplinary Methodologies, University of Warwick, UK
[4] Department of Politics and International Studies, University of Warwick, UK
[5] Getulio Vargas Foundation, Center of Public Administration and Government (FGV CEAPG), Brazil
[6] Federal University of ABC, Brazil
[7] Institute of Geography, Heidelberg University, Germany

Corresponding author:
Porto de Albuquerque, João (j.porto@warwick.ac.uk)

Current Opinion in Environmental Sustainability 2021, **49**:153–163

This review comes from a themed issue on **Transformations to sustainability: critical social science perspectives**

Edited by **Eleanor Fisher, Emily Boyd** and **Eduardo Brondizio**

For a complete overview see the Issue

Available online 24th July 2021

Received: 31 August 2020; Accepted: 18 June 2021

https://doi.org/10.1016/j.cosust.2021.06.009

Introduction

The slow pace of progress of nations around the globe to improve sustainability and tackle the United Nations' Sustainable Development Goals (SDGs) launched in 2015 has sparked calls for substantial transformations that go beyond incremental changes. The idea of transformations to sustainability has thus acquired central importance in both research [1,2] and policy discourses [3]. The devastating impacts of the COVID-19 global pandemic to economies, livelihoods and societies worldwide has given greater impetus to 'forge the transformative pathways needed to create a more livable world' [4, p. 3].

The crisis resulting from the COVID-19 pandemic has also brought about an accelerated adoption of digital technologies in many parts of the world, which have enabled many people to carry on social, economic and education activities amidst restrictions of physical contact. This positive contribution of digital technologies to strengthen the resilience of societies in face of the pandemic disruption has given rise to a renewed attention to the crucial need of 'data innovations' to support sustainability goals [4,5], echoing earlier calls for a data revolution for sustainable development [6] and the more recent emphasis on the crucial role of a 'digital revolution' to support transformations to sustainability [2,7,8•].

However, despite the acknowledged importance of digital technologies and data innovations for progress towards sustainability, there is a generalised lack of clarity on the specific transformation pathways which are to be enabled by these technological innovations, and how they are related to socio-cultural aspects, governance and politics. For instance, the investment in data innovations such as 'big data' [9,10] and citizen-generated data [11,12] is frequently justified by the need to close gaps in the data for monitoring and reporting on the SDG targets and indicators [2,4]. Although such digital innovations and emerging data sources are rightly seen as necessary for tracking and assessing progress towards the SDGs, there

is an overly optimistic underlying assumption that increased data availability will automatically lead to improved decision-making and propel transformations to more sustainable futures. Recent examples such as the deforestation of the Amazon, biodiversity loss, climate change and COVID-19 clearly show that data are important but not sufficient to compel action to change; for this, data must be presented in adequate formats for stakeholders and embedded as information into social decision-making processes with clear pathways to enable transformations.

As the critical scholarship on sustainability transformations has argued [1,13,14], there is a need to be clearer about 'what should be transformed, by and for whom, and through what processes' [14, p. 65] so as to fully account for the crucial socio-economic, governance and political aspects involved in sustainability transformations. In parallel, wider implications of datafication processes have been explored in the fields of critical data studies [15,16] and data justice [17], such as data privacy, surveillance, ownership, accessibility and inclusivity. However, such critical data issues have not been fully addressed in the sustainability discourse, in which the role of digital technologies, and data innovations in particular, have received scant attention so far in analyses of sustainability transformations.

This article attempts to fill this knowledge gap by investigating the following overarching research question: *What is the role of data innovations in transformations to sustainability*? For doing this, we first present a review of the recent literature on data-enabled transformations to sustainability, unfolding the main research question into a set of subquestions to investigate the underlying assumptions, specific roles and transformative processes enabled by data mentioned in the literature (Section 'Literature review'). The results of our review point out to crucial limitations in the way data has been conceptualised in the extant literature. In response, we introduce in Section 'Conceptual framework: transformation pathways and functions of data' a synthesis framework which provides an expanded perspective on the relationship between data and transformations to sustainability. Finally, Section 'Research agenda: exploring tensions to rethink the role of data for transformations to sustainability' presents directions for future work in this area, and Section 'Conclusion' closes the article with final remarks.

Literature review

In this paper, we understand transformations as 'fundamental shifts in human and environmental interactions and feedback' [18]. They are distinct from similar concepts widely utilised in sustainable development such as *transition* or *change* [19] due to their longer-term orientation and the gradual mainstreaming of behaviours, cultures and practices they induce [20,21]. In order to gain an understanding on the role of data innovations in such transformation processes, we took inspiration in Scoones *et al.* [14] to unfold our overarching research question into the following three sub-questions: (a) *what is being transformed with the support of data innovations?* (b) *are data innovations enabling sustainability transformations for whom and by whom?* (c) *through which processes are data innovations supporting sustainability transformations?* We then undertook a focused systematic literature review of articles published on transformations to sustainability in the past few years (2018–2020), broadly oriented by the methodological guidelines proposed by Kitchenham and Charters [22].

In order to select our primary studies, we applied the search string ('sustainability' and 'data' and 'transformation') to the electronic database SCOPUS[8] to search for studies published between January 2018 and July 2020. The choice of the relative short period of time was due to our goal of obtaining a snapshot of the recent developments in the field, whilst SCOPUS was chosen due to ease of handling and its relatively broad coverage of many scientific journals and conferences. We are conscious that the choice of keywords, timeframe and database will inevitably exclude many studies that are related to data and sustainability transformations but use other terms and indexes; however, we see the coverage of our review as a practical and meaningful sample of the most recent relevant research.

The selection of key studies was based on three inclusion criteria: (*i*) the article matches the keywords, (*ii*) the article discusses a type of transformation to sustainability and (*iii*) the article discusses how data supports transformation. As exclusion criteria, articles were discarded if they only mentioned primary/secondary data used for the study itself, but did not refer to data as part of sustainability transformation processes.

The keyword search in the Scopus database resulted in 436 primary studies. Four researchers analysed the titles and abstracts and applied the inclusion/exclusion criteria. If the purpose of the article was not clear in the abstract, three of them read it and discussed it afterwards. After this stage, 21 studies were selected for full-text reading and analysis. In the final stage of the review, each study was independently analysed in full by two members of the research team, who extracted information about our research questions.

The following three sub sections presents the findings of our review for each of our three subquestions in turn.

What is being transformed with the support of data innovations?

Table 1 presents a summary of the transformations enabled by data which are reported in each of the articles

[8] https://www.scopus.com/.

Table 1

Data-enabled sustainability transformations of the reviewed studies

References	Transformations enabled by data
Angeler *et al.* [23•]	Time series data concerning agricultural sector degradation (inundations, rainfall) are transformed into musical sounds (a 'lament') in a learning process that combines feelings and cognition (sonification), with an aspiration to induce changes in behaviours, perspectives and aims related to environmental protection.
Bibri [24•]	Simulation models based on big data analytics can be used to support decisions pertaining to optimisation, control, management, design, and planning of modern cities in an attempt to advance the contribution of smart sustainable cities to the goals of sustainable development.
Bostancı [59]	Data is being used as a means to support the application of concepts of thermodynamics in urban design through the different types of energy efficiency assessment, in an attempt to support transformations towards increasing quality of life in cities.
Cheema and Khan [25•]	ICT and IoT technologies, through the quick access to large amounts of data they provide, can support data-driven agriculture and transform existing methods, while also reducing production costs, increasing yield and profitability of farmers and improving food security, particularly in the Global South.
Dewi *et al.* [26]	Local knowledge and individual experiences have the potential to transform local communities and lead them towards a more resilient and sustainable future.
Dlugosch *et al.* [27•]	By combining data collected by different sources (sensors, IoT systems, social media) with other available authoritative data sources and through a variety of analytics and simulation methods, the potential of Shared Autonomous Electric Vehicles to minimise carbon emissions and improve urban mobility in the city of Berlin, Germany is discussed.
Dong *et al.* [28]	Big Earth data acquired through remote sensing techniques can provide invaluable information to support the search for optimal trajectories, to determine corrections and support transformations to a more sustainable future.
Dornhofer *et al.* [29]	ICT innovations, such as the green Knowledge Management Cube (KMC) architecture, are in the centre of the smart city concept and can support transformations to a more environmentally sustainable future.
Guo *et al.* [30•]	Through the establishment of a 'system of systems' for sharing environmental data and information between smaller data systems across several countries, monitoring of the state of the Earth will be enhanced.
Juneja *et al.* [31]	Big Data from various sources (i.e. sensors, smartphones) can support the process of transformation of contemporary metropolitan urban areas to smart cities.
Ketter *et al.* [32•]	Big Data can be used to build predictive models and contribute in balancing energy demand and supply, inevitably leading to transformations in energy grid management and (inter)national energy policies.
Kritzer [33•]	High quality, accurate and continuously updated data is used to enhance the management of fishery systems, with a future aspiration of improving levels of satisfaction and transforming attitudes and actions among the fishing fleets.
Pappas *et al.* [8•]	Although the application of big data analytics in businesses can improve their organisational capabilities (innovation, performance and so on) and enable digital transformations towards sustainability, the identification of different societal actors as well as the level of trust and collaboration among them needs to be thoroughly examined and analysed.
Pecora and Lins [34•]	Monitoring and management of water resources is enhanced through the introduction of an integrated Hydrological Observing System that provides a suite of visualisation tools and services to support hydrological data usage and sharing across scientific and operational communities.
Penicaud *et al.* [35]	Access to a variety of heterogeneous datasets is not enough to guarantee high quality of dairy products. A transformation in the process of organisation and editing of data from different sources and of different consistency (qualitative/quantitative) is required in order to make such data compatible to provide dairy products of higher quality.
Ratter *et al.* [36•]	Citizen perceptions collected through surveys can induce governance transformations to enhance coastal management and environmental protection in the context of climate change, as the example of Maldives demonstrate.
Romanska-Zapala *et al.* [37•]	Using the modular structure of a proposed new database capable of processing very large amounts of data, energy use is minimised while concurrently the thermal comfort of the occupants is maximised, with a potential to establish a long-term transformation in energy habits.
Saied *et al.* [38]	Big Data and other smart city applications, such as smart phones and intelligent systems, can assist in improving services provided to displaced communities and support transformation of post-war urban conglomerations to smart and sustainable cities in the long run.
Tumusiime *et al.* [39•]	By providing a diagnostic on the sources of technical and data acquisition problems for the generation of renewable energy through biogas, pathways for policy and research incentives to foster energy consumption transformation in the country are discussed.
Villegas-Ch *et al.* [60]	Big data architecture computational approach is capable of transforming not only university campus into a smart campus but also modern cities into smart cities.
Weiand *et al.* [40•]	Attitudinal data of citizens regarding their mobility patterns can inform transformation policy and improve decision-making processes.

included in our review. As this summary shows, data innovations are being used to support transformations to sustainability in relation to a wide range of application areas, such as marine ecosystems, energy efficiency, water management, urban mobility, smart cities, climate change and food production.

In order to cluster the types of transformations found in the reviewed papers and present an overview of what is being transformed through data innovations, Table 2 associates the reviewed studies with the six Sustainable Development Goal (SDG) Transformations proposed by Sachs *et al.* [2]. Whilst we have found a good coverage of

Table 2

Types of transformation identified in the literature review

SDG transformations [2]	Number of studies	References
Education and gender equality	0	N/A
Health, wellbeing and demography	1	Penicaud *et al.* [35]
Energy decarbonisation and sustainable industry	7	Romanska-Zapala *et al.* [37•], Cheema and Khan [25•], Patterson *et al.* [1], Tumusiime *et al.* [39•], Dlugosch *et al.* [27•], Weiand *et al.* [40•], Ketter *et al.* [32•]
Sustainable food, land, water and oceans	6	Kritzer [33•], Pecora and Lins [34•], Ratter *et al.* [36•], Dong *et al.* [28], Angeler *et al.* [23•], Penicaud *et al.* [35]
Sustainable cities and communities	9	Pecora and Lins [34•], Dornhofer *et al.* [29], Saied *et al.* [38], Dlugosch *et al.* [27•], Dewi *et al.* [26], Juneja *et al.* [31], Villegas-Ch *et al.* (2019), Pappas *et al.* [8•], Bibri [24•]
Digital revolution for sustainable development	4	Guo *et al.* [30•], Cheema and Khan [25•], Villegas-Ch *et al.* (2019), Pappas *et al.* [8•]

four SDG Transformations, there is a noticeable lack of studies addressing transformations on 'education and gender equality' and 'health, wellbeing and demography'.

Are data innovations enabling sustainability transformations for whom and by whom?

Our analysis has shown that the vast majority of research lays emphasis on data acquisition through different types of sensors, with most of data utilised to support top-down decision-making [29,37•]. However, a number of urban and social research on data studies, mostly apparent in the 'Sustainable Cities and Communities' transformation group, also acknowledge the importance of including citizens in decision-making processes, with data being seen as a means for citizen engagement [24•,25•].

Another group of publications emphasises that structural transformations in public policy and governance can be achieved through improved data flows across a wider array of stakeholders in different operational scales [8•,32•,34•]. This highlights the need for an increased attention from practitioners and decision-makers on improving communication channels among different groups of stakeholders: communities (micro-level) [36•,40•]; local authorities and city governments (meso-level) [26]; and national and international organisations (macro-level) [25•,30•].

Figure 1 presents an overview of these findings by visualising the connections between the socio-spatial scales and SDG Transformations which we have found in our reviewed studies. It is noticeable that the digital revolution transformations are never mentioned at the micro level, which indicates a mostly top-down view of the potential of these technologies. The largest group of studies we identified is focused on the 'sustainable cities and communities' transformation and their vast majority is situated at the meso level, which could be expected given their focus on city governance. Whilst all five SDG transformations were observed at the macro level, the micro level has received least attention in our reviewed studies.

Through which processes are data innovations supporting sustainability transformations?

Some studies we reviewed are focused on portraying the emerging capabilities presented by novel digital technologies and by the abundance of data, emphasising their *potential* for enabling transformations (e.g. Ref. [24•]), but without providing details on how this transformation potential can be realised in specific settings and contexts. In contrast, another group of studies describes *effective* transformations achieved through data innovations by presenting real-life applications of data innovations in different domains such as the monitoring of SDGs [30•], energy use [37•], food security and sustainable agriculture [25•,35] and smart cities [24•,31].

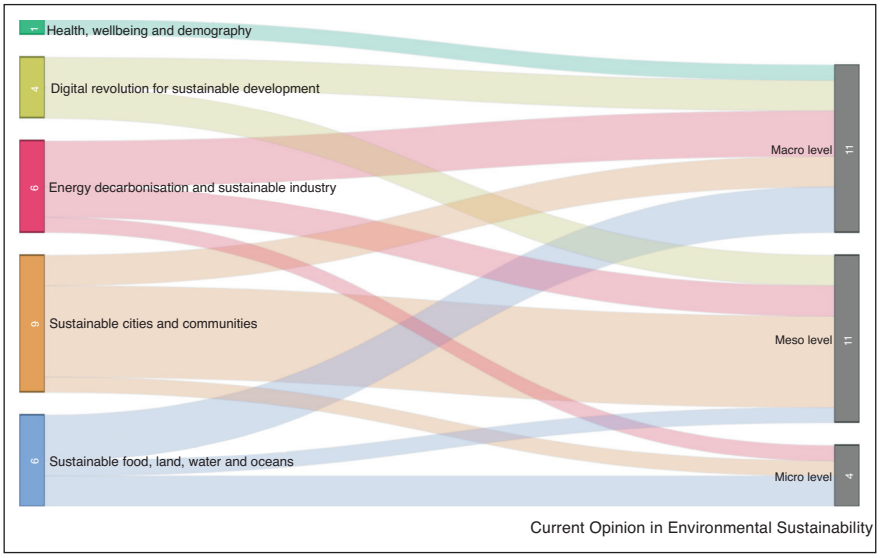
However, in most of the reviewed publications, the processes by which data innovations would engender sustainability transformations are not clearly described. In several studies, the impacts of data innovations for wider social transformation are mostly taken for granted, whilst the specific pathways with which data could enable change are not discussed. We thus conclude that a clear distinction of the pathways and roles of data to enable transformations to sustainability is an important research gap that arises from our literature review.

Conceptual framework: transformation pathways and functions of data

In response to the reviewed literature, this section introduces a synthesis framework for conceptualising the role of data innovations in transformations to sustainability. This framework attempts to systematise a set of potential pathways and functions acquired by data for enabling sustainability transformations, based both on the reviewed literature and on other related studies.

The main tenet underpinning our framework is an understanding of data not only as artefacts (i.e. binary encodings inscribed in digital media) but also as part of socio-material *processes*. We thus take a broader perspective that looks at the *data practices* through which digital artefacts are generated, transmitted, changed and used in practice, in dialogue with a growing literature on critical data studies [15,16] and data justice [17].

Figure 1



Distribution of reviewed articles per SDG Transformation type (coloured boxes on the left-hand side) and socio-spatial scale (grey boxes on the right-hand side) of communities (micro-level); local authorities and city governments (meso-level); and national and international organisations (macro-level). The numbers of reviewed articles that fall in each category are presented inside the boxes. The line thickness is scaled to the number of reviewed articles and their trajectory relate the transformation type to the corresponding socio-spatial scale. (Source: the authors).

The first component of our framework is based on the identification of the different scales, actors and types of data which can be mobilised in transformation processes, which are depicted in Figure 2. This diagram enables the recognition that actors in different scales should be considered when thinking about data-enabled sustainability transformations: (a) international/national centres of expertise in the macro level; (b) city governments and regional organisations in the meso level; (c) communities, local NGOs and other grassroots organisations in the micro level. Furthermore, it draws attention to the fact that data innovations should consider not only more traditional 'top-down' narratives (including the so-called 'big data' from centres of expertise) but also the bottom-up narratives in the form of 'thick data' generated through citizen participation processes. The diagram also draws attention to the importance of the flows between these different scales and types of data, as a means to create more robust data-enabled sustainability pathways.

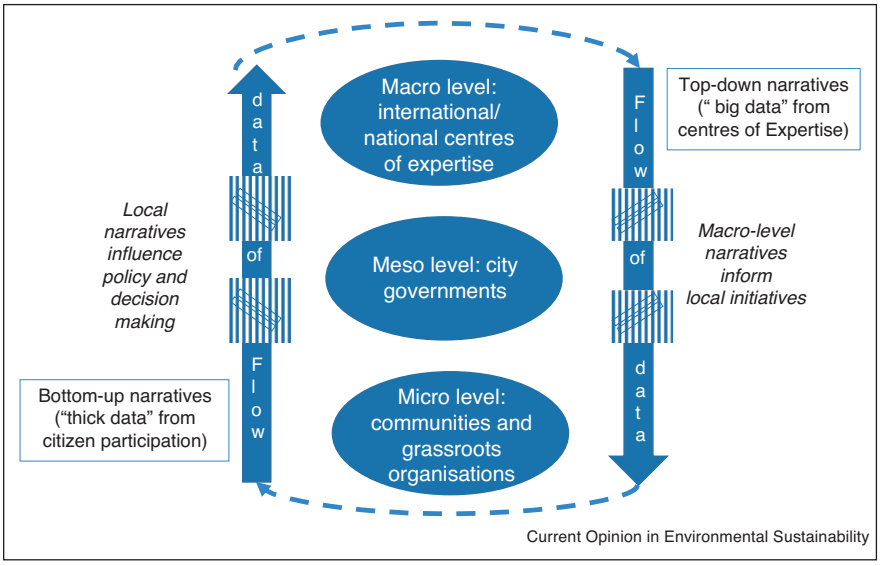
Our framework also identifies three main data-enabled transformation pathways to sustainability transformations, synthesized in Table 3: usage, circulation and generation. Within these three pathways, the role of data is modulated by what we call 'functions of data', which we identified by drawing on the classic distinction of six functions of language as proposed by the linguist Roman Jakobson [41•]: referential, metalingual, phatic, conative, expressive (also called emotive) and poetic (see Table 3). These functions were proposed to provide a broader view on the different functions acquired by language in the pragmatics of actual speech events. Analogously, we employ the functions of data within our framework to enable more specific accounts on the role acquired by data in actual practices of production, circulation and usage of digital artefacts in data innovations. Table 3 provides examples of studies in our literature review in which we could identify a main reference to each pathway and function of data. However, it is important to notice that these pathways and functions are not mutually exclusive; several research studies and practical projects will address a combination of the pathways and functions of the proposed framework.

We explain each of the data-enabled transformation pathways and functions of data of our framework in the next sections.

Data usage

The first and most evident pathway for data to enable transformations to sustainability is focused on data usage: making sense of data can inform decisions and actions supporting transformations to sustainability. Here the most important function of data (similarly to language) is to establish a reference to the context of transformations, that is, data acquires a *referential* function by indexing or measuring a particular contextual element, which thus serves as evidence to inform decision-making related to transformations to sustainability (e.g. Refs. [43–45,58]). The prevalence of this function is confirmed by our analysis of the reviewed studies, all of which do include a referential function in one way or another: for instance,

Figure 2



Scales, actors and types of data which can be mobilised in sustainability transformations (Source: the authors).

data has been used as indicators to monitor the Sustainable Development Goals [30•], to track energy usage [37•]; to quantify environmental impacts of food systems [25•,39•] and to represent socio-ecological-technological systems in support of future planning for sustainability transformations [61]. The data usage transformation pathway is also the most widely referred to in policy discourses around data innovations mentioned in the introduction [2,4].

Nevertheless, data may have additional functions in practice which are concomitant to its main referential role. Data, for instance, can enact a *poetic* function when its aesthetic affordances are brought to the fore by means of creative data visualisations [46] or by 'reading' data through physical devices to provoke embodied sensations in what Calvillo and Garnett [47] call 'data intimacies'. Among our reviewed studies, Angeler *et al.* [23•] present a good example of a poetic data function: a 'sonification'

Table 3

Pathways and functions of data in transformations to sustainability

Data-enabled transformation pathways	Functions of data	Example studies
Data usage: the interpretation of data provides evidence to support decision-making, informs desired/intended transformations and enables monitoring and tracking progress.	**Referential**: data offers indexical access to contextual elements to support transformations.	Guo *et al.* [30•], Cheema and Khan [25•], Romanska-Zapala *et al.* [37•], Dlugosch *et al.* [27•], Tumusiime *et al.* [39•], Dong et al [28], Dornhofer *et al.* [29], Kritzer [33•], Penicaud *et al.* [35], Bibri [24•], Juneja *et al.* [31]
	Poetic: the form and aesthetic affordances of data can be explored as a poetic manifestation.	Angeler *et al.* [23•]
Data circulation: the flow of data between different actors and scales can enable recognition and coordination, facilitating change in governance arrangements and opening up new communication channels.	**Phatic**: data exchange creates and maintains communication channels by allowing recognition of actors.	Pappas *et al.* [8•], Dewi *et al.* [26], Weiand *et al.* [40•]
	Conative: circulation of data engages a receptor/stakeholder by addressing them explicitly and building trust.	Ratter *et al.* [36•]
Data generation: creating new data is a transformative opportunity in itself as a catalyst for mutual social learning, development of critical consciousness and change of perspectives and behaviours.	**Metalingual**: data enables reflection about the issues represented, social processes, formats/ standards and worldviews.	Pecora and Lins [34•]
	Expressive: data production can be an expressive medium to give voice to hitherto invisible personal and emotive connections to social and environmental phenomena.	Garde-Hansen *et al.* [42•]

method converts a data time series about inundation and rainfall into musical sounds. The resulting music (called a 'lament') is used to testify environmental degradation, thus opening up opportunities to transform community behaviours through engagement and action [23•].

Data circulation

The flow of data between different actors and systems offers an additional transformation pathway for data, which can enable coordination and communication between hitherto disconnected actors. Independent of the references codified in data, the mere act of exchanging data may have a *phatic* function, that is, it can serve as a platform for actors to create and maintain communication channels, engendering new understandings and perspectives on sustainability issues and enabling changes in governance arrangements and organisational structures [8•,26,40•,57].

Furthermore, in analogy to the use of imperatives in language, data can be used with a *conative* function to address and engage specific groups of actors with a request which may in turn support transformations. For instance, the process of circulating data from citizens to government (and vice-versa) can support transformative governance not only by increasing the diversity of perceptions taken into account into decision-making (i.e. having a referential function) but also as a result of stronger engagement of hitherto excluded interest groups, leading to trust building, broader consensus and wider support for a climate-resilient sustainable development pathway [36•].

Data generation

A third data-enabled transformation pathway can be distinguished by regarding the data generation process as a transformative moment in itself, in addition to the potential transformations that the data contents can inform in the future, or to the effects emerging from data circulation. For instance, the process of generating data with citizens can be leveraged as an opportunity for social learning, empowering disenfranchised stakeholders and enabling a new critical consciousness about the sustainability issues which are intended to be captured with data [48•,49•].

In our framework, data generation is associated with a *metalingual* function, that is, data is able to refer not only to an external context, but also to its own technical formats, categories, coding schemes and supporting media, thus working as meta-data. A practical example of exploiting the metalingual function to enable transformations is found in the work of Pecora and Lins [34•], which describes how hydrological meta-data, the specification of data formats and an ontology (i.e. a logical data scheme) are able to help clarify the meanings of available data for all involved. This process can support the creation of a shared understanding about which data are useful for supporting decisions and thereby it can stimulate the generation of missing data.

Although we could not find any corresponding study in our review, we added one last function of data to our framework based on Jakobson [41•] and a related study [42•]. The *expressive* or *emotive* function comes from the fact that the informational capacities of data are not restricted to its contents. The way with which a sentence is uttered (including pauses, emphases, interjections) can be varied to express different emotions and such *expressive* function of language is an integral part of its informational capacities. Analogously, data generation can be used as an expressive medium to empower hitherto invisible social groups to voice their personal and collective emotive connections to social and environmental phenomena, which could create a powerful pathway to change. For instance, data generation can surface encounters with environmental phenomena such as floods, so that the production of such digital flood memories can be used to enable change of perceptions and behaviours towards improving community resilience [42•].

Research agenda: exploring tensions to rethink the role of data for transformations to sustainability

This section outlines a critical research agenda which builds upon the literature review and conceptual framework to propose a set of challenges and critical questions to be investigated in future studies for advancing our understanding on the role of data innovations in sustainability transformations.

Table 4 presents the three data-enabled transformation pathways of our conceptual framework, and acknowledges that each of these pathways is related to corresponding challenges and risks of 'side effects', which are frequently overlooked in discourses around data. By tensioning each potential data transformation pathway against its corresponding risks and challenges, we thus derived a set of critical questions which should be reflected upon and addressed in future research and practice, summarised as follows.

Data generation

Transformative pathways related to data generation are the ones which received less attention in the literature so far. Most of the literature seems to consider the production of data as a 'means to end', and this includes some papers focusing on data generated by communities and citizens. Building upon recent recognition of the significance of citizen science [12,50,11], local and indigenous knowledge [51] in the agenda of sustainability, future research should investigate ways to leverage the potential transformative pathways and functions of data generation that we identified in our framework for achieving mutual

Table 4

Critical research agenda for the role of data in transformations to sustainability

Data-enabled transformation pathways	Challenges	Critical research questions
Data **generation** as a catalyst of transformations on perspectives and behaviours, empowering communities and creating critical consciousness	Instrumentality of citizen sensing	Who defines which data is being produced and how? Is the data generation building new capacities and critical consciousness or contributing to reduce inequalities?
	Inability to challenge bias and discrimination	Do the processes of data generation empower to contest and challenge existing assumptions?
Data **circulation** facilitates changes in governance arrangements	Data surveillance, data privacy	What are the implications of being datafied for human rights of individuals?
	Power asymmetries, Sharing in data's benefits	How are the benefits of data circulation distributed by the different actors involved? Who controls/defines the flows and methods with which data are processed and visualised? Who is being addressed by data and how?
Data **usage** informs decision-making for transformations	Representation bias, aggregation, ecological fallacies, missing data	Who is represented in the data and who is left out? How are missing data and uncertainties acknowledged, communicated and interpreted?
	Establishment of regimes of perceptibility, power asymmetries, epistemic justice	Who defines what counts as data and which data is important? Whose voices and worldviews are shaping the methods and tools used for data generation and usage? What are the social and material processes for building trust in data and how this shapes decision-making in practice?

learning and empowerment. However, it will be crucial to consider challenges arising from the way with which citizens are engaged in these processes: in order to overcome the risk of citizens being instrumentalised to gather data which is only relevant for others [48•], critical questions need to be asked, such as who defines which data is collected, and whether the processes are truly open to contestation and challenge of bias and discrimination embedded in current digital platforms and tools [17].

Data circulation

A small number of studies we reviewed acknowledged the potential of the circulation of data as a transformation pathway, but we believe that future work can further explore the transformative role of data through its capacity to facilitate change in governance arrangements, to create and maintain new communication channels among stakeholders as well as to engage-specific social actors. Nevertheless, important challenges also arise from the flow of data: for instance, the ever-increasing usage of digital means for communication makes it easy to gather massive amounts of data for surveillance [52], raising questions on what are the implications of 'datafication' for the basic human rights to privacy and freedom. It will also be imperative to ask who benefits from data flows, as existing structural power asymmetries may lead to what has been called 'technocolonialism' [53]: an unequal distribution of benefits arising from data which disfavour marginalised groups.

Data usage

We found our reviewed studies to be apace with current policy discourses in the acknowledgement of the potential transformation pathways opened up by data innovations through the usage of data to inform decision-making processes. We concur with these studies, but Table 4 also draws attention to important challenges which are associated with the referential function of data, such as biases, missing data, and privacy. This leads to important critical questions on who is represented (or not) in the data, but also on who defines what counts as data: data not only makes some things visible, but also necessarily hides other things [54•] thus defining what is called by Murphy [62] 'regimes of perceptibility'.

Furthermore, we emphasise the need for research to capture a more detailed understanding of the 'social life of data' [55], that is, to pay attention to the socio-material practices with which data 'intervene' in the unfolding of actual decision-making processes [56]. This is an important point, since an overemphasis on generation of data without fully considering their integration into social processes of trust-building and decision-making, may leave the putative transformative potential of data completely unrealised. Even worse, as tragically illustrated by the recent trends of 'fake news' and the 'post-truth' related to data on climate change, rainforest deforestation and the COVID-19 pandemic, not only missing data, but also their systematic denial by decision makers and misinformation campaigns can have tragic consequences for the environment, human health and wellbeing. Therefore, it is imperative for future research to develop more nuanced understandings of the role of data usage in practice.

Conclusion

The review, conceptual framework and critical research agenda laid down in this article offer a starting point to

develop a deeper understanding of the relation between data innovations and transformations to sustainability which includes broader considerations of social, cultural and political issues. Building upon our proposed framework and critical agenda, future research should investigate how specific projects on sustainability transformations are able to integrate the different data transformation pathways and functions we identified, whilst addressing the corresponding tensions and challenges. We thus hope this article can inspire future research and practical projects to consciously reflect about their assumptions and practices to be able to effectively integrate data into sustainability transformation processes.

Data access statement

All data is provided in full in the results section of this paper.

Conflict of interest statement

Nothing declared.

Acknowledgements

The authors are grateful to Philipp Ulbrich for his contributions to the diagram of Figure 2.

This article is part of the project Waterproofing Data that is financially supported by the Belmont Forum and NORFACE Joint Research Programme on Transformations to Sustainability (https://www.norface.net/program/transformations-to-sustainability/), which is co-funded by DLR/BMBF, ESRC/GCRF (ES/S006982/1), FAPESP (process n. 18/50039-4) and the European Commission through Horizon 2020. Fernanda Lima Silva and Mario Henrique da Mata Martins acknowledge support from the São Paulo State Research Support Foundation (FAPESP) for their postdoctoral fellowships (project numbers 2019/06616-0 and 2019/06595-2). Lívia Castro Degrossi acknowledges the funding by the UK Research and Innovation through the GCRF Global Research Translation Award (EPSRC grant: EP/T015683/1).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest

1. Patterson J, Schulz K, Vervoort J, van der Hel S, Widerberg O, Adler C, Hurlbert M, Anderton K, Sethi M, Barau A: **Exploring the governance and politics of transformations towards sustainability**. *Environ Innov Soc Transit* 2017, **24**:1-16 http://dx.doi.org/10.1016/j.eist.2016.09.001.

2. Sachs JD, Schmidt-Traub G, Mazzucato M, Messner D, Nakicenovic N, Rockström J: **Six transformations to achieve the sustainable development goals**. *Nat Sustain* 2019, **2**:805-814 http://dx.doi.org/10.1038/s41893-019-0352-9.

3. UN-DESA: **Sustainable development outlook 2019**. *Sustainable Development Outlook 2020*. United Nations Department of Economic and Social Affairs; 2020 http://dx.doi.org/10.18356/7a3ee84a-en.

4. UN: *The Sustainable Development Goals Report 2020*. United Nations; 2020 http://dx.doi.org/10.29171/azu_acku_pamphlet_k3240_s878_2016.

5. Sachs JD, Schmidt-Traub G, Kroll C, Lafortune G, Fuller G, Woelm F: **The sustainable development goals and COVID-19**. *Sustainable Development Report 2020*. Cambridge University Press; 2020.

6. UN: *A World that Counts: Mobilising the Data Revolution for Sustainable Development*. United Nations Data Revolution Task Force; 2014.

7. Corbett J, Mellouli S: **Winning the SDG battle in cities: how an integrated information ecosystem can contribute to the achievement of the 2030 sustainable development goals**. *Inf Syst J* 2017, **27**:427-461 http://dx.doi.org/10.1111/isj.12138.

8. • Pappas IO, Mikalef P, Giannakos MN, Krogstie J, Lekakos G: **Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies**. *Inf Syst E-Bus Manag* 2018, **16**:479-491 http://dx.doi.org/10.1007/s10257-018-0377-z
The article is part of a special edition on big data and business analytics ecosystems and proposes a model for the digital transformation and promotion of sustainable cities based on the integration of multiple actors, understanding of their interrelations and potential capabilities for big data analytics.

9. Kitchin R: **Conceptualising data**. *The Data Revolution: Big Data, Open Data, Data Infrastructures and their Consequences*. Sage; 2014:1-26.

10. Mayer-Schönberger V, Cukier K: *Big Data: A Revolution that Will Transform how we Live, Work, and Think*. Houghton Mifflin Harcourt; 2013 http://dx.doi.org/10.2501/ija-33-1-181-183.

11. Läummerhirt D, Gray J, Venturini T, Meunier A: *Advancing Sustainability Together? Citizen-Generated Data and the Sustainable Development Goals*. Global Partnership for Sustainable Development Data; 2018 http://dx.doi.org/10.2139/ssrn.33204678.

12. Fraisl D, Campbell J, See L, Wehn U, Wardlaw J, Gold M, Moorthy I, Arias R, Piera J, Oliver JL *et al.*: **Mapping citizen science contributions to the UN sustainable development goals**. *Sustain Sci* 2020, **15**:1735-1751 http://dx.doi.org/10.1007/s11625-020-00833-7 0123456789.

13. Feola G: **Societal transformation in response to global environmental change: a review of emerging concepts**. *Ambio* 2015, **44**:376-390 http://dx.doi.org/10.1007/s13280-014-0582-z.

14. Scoones I, Stirling A, Abrol D, Atela J, Charli-joseph L, Eakin H, Ely A, Olsson P, Pereira L, Priya R, Van Zwanenberg P: **Transformations to sustainability: combining structural, systemic and enabling approaches**. *Curr Opin Environ Sustain* 2020, **42**:65-75 http://dx.doi.org/10.1016/j.cosust.2019.12.004.

15. Dalton C, Thatcher J: **What does a critical data studies look like, and why do we care? Seven points for a critical approach to 'big data'**. *Soc Space* 2014, **29**.

16. Schäfer MT, van Es K: **The datafied society. Studying culture through data**. In *The Datafied Society. Studying Culture through Data.* Edited by Tobias Schäfer ro, van Es K. Amsterdam University Press; 2017 http://dx.doi.org/10.5117/9789462981362.

17. Taylor L: **What is data justice? The case for connecting digital rights and freedoms globally**. *Big Data Soc* 2017, **4**:1-14 http://dx.doi.org/10.1177/2053951717736335.

18. Hölscher K, Wittmayer M, Loorbach D: **Transition versus transformation: what's the difference?** *Environ Innov Soc Transit* 2018, **27**:1-3.

19. Loorbach D, Rotmans J: **Managing transitions for sustainable development**. In *Understanding Industrial Transformation. Environment & Policy*, , vol 44. Edited by Olsthoorn X, Wieczorek A. Dordrecht: Springer; 2006 http://dx.doi.org/10.1007/1-4020-4418-6_10.

20. Bahadur A, Tanner T: **Transformational resilience thinking: putting people, power and politics at the heart of urban climate resilience**. *Environ Urban* 2014, **26**:200-214 http://dx.doi.org/10.1177/095624781452215.

21. Pitidis V, Coaffee J: **Catalysing governance transformations through urban resilience implementation: the case of Thessaloniki, Greece**. *Cities* 2020, **107** http://dx.doi.org/10.1016/j.cities.2020.102934.

22. Kitchenham BA, Charters S: *Guidelines for Performing Systematic Literature Reviews in Software Engineering*. 2007.

23. Angeler DG, Alvarez-Cobelas M, Sánchez-Carrillo S: **Sonifying social-ecological change: a wetland laments agricultural transformation**. *Ecol Soc* 2018, **23**:1-13 http://dx.doi.org/10.5751/ES-10055-230220
In this paper, the authors show how the circulation of ecosystem data in audio format can potentially raise awareness about the unsustainable use of water supplies and other socio-ecological challenges and thus support changes in perception and behaviours.

24. Bibri SE: **Unprecedented innovations in sustainable urban planning: novel analytical solutions and data-driven decision-making processes**. *Urban Book Series*. 2018 http://dx.doi.org/10.1007/978-3-319-73981-6_5
The book chapter gives a detailed overview of the potential of big data analytics for smart sustainable urban planning. The chapter includes information about various related aspects such as urban planning, sustainable smart cities, and sources of urban big data. Challenges and opportunities are portrayed. Overall, it is concluded that data-driven decision making and big data can be of great value for planning sustainable development in urban areas.

25. Cheema MJM, Khan MA: **Information technology for sustainable agriculture**. *Innovations in Sustainable Agriculture*. 2019 http://dx.doi.org/10.1007/978-3-030-23169-9_19
The main argument lies around the potential of ICT and IoT technologies and of data-driven agriculture in general to reduce production costs and increase yield and profitability of farmers while simultaneously providing better food security, particularly in the Global South. Real-time Data Analytics are particularly endorsed as a means of supporting local farmers and leading to Precision Agriculture ultimately maximising both the environmental and local farmers' benefits.

26. Dewi AE, Maryono M, Warsito B: **Potential sustainability of "Kampung Iklim" program in Surakarta municipal**. *E3S Web of Conferences* 2019, **vol 125** http://dx.doi.org/10.1051/e3sconf/201912502002.

27. Dlugosch O, Brandt T, Neumann D: **Combining analytics and simulation methods to assess the impact of shared, autonomous electric vehicles on sustainable urban mobility**. *Inform Manag* 2020 http://dx.doi.org/10.1016/j.im.2020.103285
The article conducts a case study on urban mobility in the city of Berlin with a focus on the potential impacts of shared autonomous electric vehicles. It concludes that this type of transport has the potential to substantially reduce investments in resources and that sharing this information with different spheres can support the promotion of sustainable urban mobility.

28. Dong J, Metternicht G, Hostert P, Fensholt R, Chowdhury RR: **Remote sensing and geospatial technologies in support of a normative land system science: status and prospects**. *Curr Opin Environ Sustain* 2019, **38**:44-52 http://dx.doi.org/10.1016/j.cosust.2019.05.003.

29. Dornhofer M, Weber C, Zenkert J, Fathi M: **A data-driven smart city transformation model utilizing the green knowledge management cube**. *5th IEEE International Smart Cities Conference, ISC2 2019* 2019:691-696 http://dx.doi.org/10.1109/ISC246665.2019.9071703.

30. Guo H, Nativi S, Liang D, Craglia M, Wang L, Schade S, Corban C, He G, Pesaresi M, Li J *et al.*: **Big Earth Data science: an information framework for a sustainable planet**. *Int J Digit Earth* 2020, **13**:743-767 http://dx.doi.org/10.1080/17538947.2020.1743785
The authors propose a new engineering discipline called Big Earth Data Science, focusing on the development of theories and knowledge about the evolution of Earth's socio-physical systems. It integrates data sensing and observational data and offers a useful technological framework.

31. Juneja S, Juneja A, Anand R: **Role of big data as a tool for improving sustainability for the betterment of quality of life in metro cities**. *Int J Control Autom* 2019, **12**:553-557.

32. Ketter W, Collins J, Saar-Tsechansky M, Marom O: **Information systems for a smart electricity grid: emerging challenges and opportunities**. *ACM Trans Inf Syst* 2018, **9** http://dx.doi.org/10.1145/3230712
The paper is centered on examining the Green IS with a focus on sustainable energy (Smart Grid); in particular, the emergence of sustainable and economically viable energy systems that integrate energy production, distribution, storage and consumption. It is a secondary study that reviews existing literature and summarizes key challenges and relevant ongoing research.

33. Kritzer JP: **Influences of at-sea fishery monitoring on science, management, and fleet dynamics**. *Aquacult Fish* 2020, **5**:107-112 http://dx.doi.org/10.1016/j.aaf.2019.11.005
In order to overcome overexploitation processes, the author proposes that effective at-sea monitoring need to be implemented with the aim of achieving sustainability.

34. Pecora S, Lins HF: **E-monitoring the nature of water**. *Hydrol Sci J* 2020, **65**:683-698 http://dx.doi.org/10.1080/02626667.2020.1724296
The authors suggest WMO Hydrological Observing System operational architecture is hydrology's system for the future because it addresses societal needs for multiple environmental areas, such as disaster risk reduction, sustainable resources, climate resilience and economic growth.

35. Penicaud C, Ibanescu L, Allard T, Fonseca F, Dervaux S, Perret B, Guillemin H, Buchin S, Salles C, Dibie J, Guichard E: **Relating transformation process, eco-design, composition and sensory quality in cheeses using PO2 ontology**. *Int Dairy J* 2019, **92**:1-10.

36. Ratter B, Hennig A, Zahid: **Challenges for shared responsibility – political and social framing of coastal protection transformation in the Maldives**. *J Geogr Soc Berlin* 2019, **150**:169-183 http://dx.doi.org/10.12854/erde-2019-426
The authors challenge the assumption that people from Maldives prefer hard coastal protection since national politicians have actually excluded them from governance processes regarding these issues.

37. Romanska-Zapala A, Dudek P, Górny M, Dudzik M: **Modular statistical system for an integrated environmental control**. *E3S Web of Conferences* 2020, **172** http://dx.doi.org/10.1051/e3sconf/202017219006
This paper describes the Modular Statistical Software (MSS) that performs data acquisition, processing and analysis of observations related to indoor weather in buildings. The information is transferred to the buildings automated control (BAC) that operates the heating and cooling systems to maintain indoor thermal comfort with energy use efficiency.

38. Saied Y, Rodionovskaya I, Nassour M, Husen R: **The role of smart city applications in the development of IDP areas in war countries**. *IFAC-PapersOnLine* 2019, **52**:246-251 http://dx.doi.org/10.1016/j.ifacol.2019.12.490.

39. Tumusiime E, Kirabira JB, Musinguzi WB: **Long-life performance of biogas systems for productive applications: the role of R&D and policy**. *Energy Rep* 2019, **5**:579-583 http://dx.doi.org/10.1016/j.egyr.2019.05.002
The authors start from a case study in Uganda to argue that an approach focused on research and development, with major focus on data, corroborates to achieve the success of productive biogas installations and is therefore necessary for renewable energy policy development.

40. Weiand L, Schmitz S, Becker S, Niehoff N, Schwartzbach F, Von Schneidemesser E: **Climate change and air pollution: the connection between traffic intervention policies and public acceptance in a local context**. *Environ Res Lett* 2019, **14**:85008 http://dx.doi.org/10.1088/1748-9326/ab299b
In the context of the implementation of a traffic measure, this article characterizes groups of citizens describing their mobility habits, their attitudes towards the measure, and level of environmental concern. They present group-specific barriers and opportunities, as well as pathways to encourage more sustainable transportation use.

41. Jakobson R: **Linguistics and poetics**. *Style in Language*. MIT Press; 1960:350-377
The author offers a framework for classifying six language functions (emotional, conative, referential, poetic, factual and metalingual) whose analogy with data studies may be a relevant analytical matrix to understand the different functions of the data.

42. Garde-Hansen J, McEwen L, Holmes A, Jones O: **Sustainable flood memory: remembering as resilience**. *Mem Stud* 2017, **10**:384-405 http://dx.doi.org/10.1177/1750698016667453
This article shows how memories of floods can be used as emotive and poetic expressions of lived experiences, which can be mobilised to improve community resilience to flooding.

43. Park Jongsoon, Cullen Ralph, Smith-Jackson Tonya: **Designing a decision support system for disaster management and recovery**. *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting* 2014.

44. Restrepo-Estrada Camilo, de Andrade Sidgley, Abe Narumi, Fava Maria, Mendiondo Eduardo, De Albuquerque Joao: **Geo-social media as a proxy for hydrometeorological data for streamflow estimation and to improve flood monitoring**. *Comput Geosci* 2017, **111**:148-158 http://dx.doi.org/10.1016/j.cageo.2017.10.010.

45. Danubianu Mirela: **Big data – catalyst of sustainable development**. *Pres Environ Sustain Dev* 2020, **14** http://dx.doi.org/10.15551/pesd2020141024.

46. McInerny G: **Visualizing data: a view from design space**. In *Routledge Handbook of Interdisciplinary Research Methods.* Edited by Lury C, Fensham R, Heller-Nicholas A, Lammes S, Last A, Michael M, Uprichard E. Routledge; 2018:133-141 http://dx.doi.org/10.1787/eee82e6e-en. Issue 2017.

47. Calvillo N, Garnett E: **Data intimacies: building infrastructures for intensified embodied encounters with air pollution**. *Sociol Rev* 2019, **67**:340-356 http://dx.doi.org/10.1177/0038026119830575.

48. • Albuquerque JP, Almeida AA: **Mo des of engagement: reframing 'sensing' and data generation in citizen science for empowering relationships**. In *Toxic Truths: Environmental Justice and Citizen Science in a Post Truth Age.* Edited by Davies T, Mah A. Manchester University Press; 2020

This paper proposes a view of data generation processes that draws on the critical pedagogy of Paulo Freire to investigate the need for dialogical processes of data production.

49. • Trajber R, Walker C, Marchezini V, Kraftl P, Olivato D, Hadfield-Hill S, Zara C, Fernandes Monteiro S: **Promoting climate change transformation with young people in Brazil: participatory action research through a looping approach**. *Action Res* 2019, **17**:87-107 http://dx.doi.org/10.1177/1476750319829202

This article articulates the theoretical framings of citizen science and nexus thinking under a framework of participatory action research, including data generation in a process of mutual learning in relation to the knowledge, action and critique co-produced with young people.

50. de Sherbinin A, Bowser A, Chuang T, Cooper C, Danielsen F, Edmunds R, Elias P, Faustman E, Hultquist C, Mondardini R *et al.*: **The critical importance of citizen science data**. *Front Clim* 2021, **3**:1-7 http://dx.doi.org/10.3389/fclim.2021.650760.

51. Mastrángelo ME, Pérez-Harguindeguy N, Enrico L, Bennett E, Lavorel S, Cumming GS, Abeygunawardane D, Amarilla LD, Burkhard B, Egoh BN *et al.*: **Key knowledge gaps to achieve global sustainability goals**. *Nat Sustain* 2019, **2**:1115-1121 http://dx.doi.org/10.1038/s41893-019-0412-1.

52. Zuboff S: **Big other: surveillance capitalism and the prospects of an information civilization**. *J Inf Technol* 2015, **30**:75-89 http://dx.doi.org/10.1057/jit.2015.5.

53. Madianou M: **Technocolonialism: digital innovation and data practices in the humanitarian response to refugee crises**. *Soc Media Soc* 2019, **5** http://dx.doi.org/10.1177/2056305119863146.

54. • Calvillo N: **Political airs: from monitoring to attuned sensing air pollution**. *Soc Stud Sci* 2018, **48**:372-388 http://dx.doi.org/10.1177/0306312718784656

This paper provides a theoretical lens to 'attuned sensing' and examines the case of air pollution monitoring in Madrid, going beyond a mere capture of environmental variable to understand how monitoring redistributed actors, practices and objects that make the toxicity not only knowable, but also accountable, and most importantly, opened up spaces for citizen intervention.

55. Bates J, Lin Y-W, Goodale P: **Data journeys: capturing the socio-material constitution of data objects and flows**. *Big Data Soc* 2016, **3**:205395171665450 http://dx.doi.org/10.1177/2053951716654502.

56. Tkacz N, Henrique da Mata Martins M, Porto de Albuquerque J, Horita F, Dolif Neto G: **Data diaries: a situated approach to the study of data**. *Big Data Soc* 2021, **8** http://dx.doi.org/10.1177/2053951721996036.

57. Piattoeva N, Centeno VG, Suominen O, Rinne R: **Governance by data circulation? The production, availability, and use of national large-scale assessment data**. In *Politics of Quality in Education: A Comparative Study of Brazil, China, and Russia.* Edited by Kauko ak, Rinne so, Takala oa. Routledge; 2018.

58. Horita FEA, Porto de Albuquerque J, Marchezini V, Mendiondo EMA: **Bridging the gap between decision-making and emerging big data sources: an application of a model-based framework to disaster management in Brazil**. *Decis Support Syst* 2017, **97**:12-22 http://dx.doi.org/10.1016/j.dss.2017.03.001.

59. Bostanci S: **Critical thinking about urban studies linked with thermodynamic terms**. *IOP Conf Ser Mater Sci Eng* 2019, **618** http://dx.doi.org/10.1088/1757-899x/618/1/012078.

60. Villegas-Ch W, Palacios-Pacheco X, Luján-Mora S: **Application of a smart city model to a traditional university campus with a big data architecture: a sustainable smart campus**. *Sustainability* 2019, **11** http://dx.doi.org/10.3390/su11102857.

61. Egerer M, Haase D, Frantzeskaki N, Andersson E: **Urban change as an untapped opportunity for climate adaptation**. *Nat Urban Sustain* 2021, **22** http://dx.doi.org/10.1038/s42949-021-00024-y.

62. Murphy M: *Sick Building Syndrome and the Problem of Uncertainty*. Duke University Press; 2006.