Patterns

Citizen science as a data-based practice: A consideration of data justice

Highlights

- A "data justice" analytical framework was used to study citizen science

- Five dimensions of data justice were explored in citizen science cases

- Some forms of injustice were found in citizen science cases under review

Authors

Debora Irene Christine, Mamello Thinyane

Correspondence

debora@unu.edu (D.I.C.), mamello@unu.edu (M.T.)

In brief

Despite the burgeoning field of citizen science as a data-based practice, the structural issues of the data processes in citizen science remain understudied. To address this gap, this article uses a "data justice" framework to explore social injustices emanating from the data processes in citizen science cases and considers their exploitative and exclusionary implications on citizens. As such, this work offers critical examination of citizen science and contributes to the scholarship of citizen science and critical data studies.

CellPress

Article

Citizen science as a data-based practice: A consideration of data justice

Debora Irene Christine[1,2,*] and Mamello Thinyane[1,*]

[1]United Nations University Institute in Macau, Casa Silva Mendes, Estrada do Engenheiro Trigo No. 4, Macau SAR, China
[2]Lead contact
*Correspondence: debora@unu.edu (D.I.C.), mamello@unu.edu (M.T.)
https://doi.org/10.1016/j.patter.2021.100224

THE BIGGER PICTURE   Citizen science has been regarded for its contribution to scientific research, inclusive science engagement, and addressing of social justice issues. Within citizen science, social justice is pursued through different approaches, including facilitating public participation in research and utilizing citizen science data in social justice advocacy. Although citizen science is a data-based practice, the structural dimensions of the data processes that support and hamper the pursuit of social justice in citizen science remain understudied. This article applies a "data justice" framework to unpack the elements and practices that constitute the generation, circulation, and use of data and data-related outcomes in citizen science. We demonstrate the relevance and limitations of the framework with regard to the domain of citizen science. This work thus contributes to the growing research interest in critical data studies, i.e., the study around equity issues in data science.

1 2 3 4 5   **Concept**

SUMMARY

Citizen science has been motivated by several perspectives, including increased efficiency in data collection and distributed analysis, democratizing knowledge production, making science more responsive to community needs, and improving the representation of marginalized populations in public data. Despite the potential of citizen science to achieve social justice agendas through a data-intensive and data-driven participatory scientific enquiry, scholarship in critical data studies offers several problematizations of data-based practices, highlighting risks of exclusion and inequality. To understand the extent to which citizen science supports and challenges forms of injustice, this study used a "data justice" analytical framework to critically explore the assemblages of citizen science. We examined four citizen science cases with different levels of citizen engagement, intended outcomes, and data systems. The analysis suggests instances of injustice occurring throughout the data processes of the citizen science cases across the dimensions of procedural, instrumental, rights-based, structural, and distributive data justice.

INTRODUCTION

Citizen science is a form of research collaboration that enlists the public in scientific research to address real-world problems.[1] The objectives of citizen science include producing and disseminating scientific knowledge and broadening participation in science itself.

In citizen science, citizens are engaged in generating, preparing, and processing empirical observations and detailed measurements, which are traditionally performed by professional researchers or scientists. As it gathers data and generates

knowledge through processes outside the mainstream scientific epistemology, it is a form of knowledge production wherein "communities or networks of citizens … act as observers in some domain of science."[2]

Citizen science has been motivated by several perspectives, such as the instrumental goal of leveraging the engagement of citizens for increased efficiency in data collection and distributed analysis. It has further been expanding as an avenue for making science more responsive to community needs by facilitating the broadening and deepening of the engagement of underrepresented groups in some aspect of scientific research, as well as

Table 1. Types and levels of public participation in science[58,66]

Citizen science model	Description of interactions between professional researchers and public participants	Participation dimension
Contractual	Communities ask professional researchers to conduct a specific scientific investigation and report on the results	Nominal
Contributory	Projects are generally designed by scientists and members of the public primarily contribute data to them	
Collaborative	Projects are generally designed by scientists and members of the public contribute data but also help to refine project design, analyze data, and/or disseminate findings	Instrumental
Co-created	Projects are generally designed by scientists and members of the public working together and at least some of the public participants are actively involved in most or all aspects of the research process	Representative
Collegial	Non-credentialed individuals conduct research independently with varying degrees of expected recognition by institutionalized science and/or professionals	Transformative

acquiring visibility, in and through public data, to remedy specific social injustices. Citizen science also has strong motivation from the position of rights, in that the Universal Declaration of Human Rights recognizes the right of everyone "to share in scientific advancement and its benefits."[3] An example outworking of this human rights position was the Aarhus Convention, which codified the rights of citizens to participate in decision-making and have access to justice on projects affecting the environment.[4]

Traditionally, citizen science projects have focused on the natural sciences, such as natural resources management,[5] environmental monitoring and preservation,[6] and astrophysics.[7] Today, there is a wide range of models of public engagement in scientific research, distinguished from one another based on the extent to which the public is involved in the scientific research process and the type of its contribution to the research; well-established variations include participatory research,[8] community-based monitoring,[9] and crowdsourced science,[10] among others.

The burgeoning of citizen science is accompanied by increasing theorizing of what is considered science ("boundary-work")[11] and what is considered knowledge[12] and participation in the field of citizen science.[1] The diverse forms of citizen science are related to the development of several research methodologies, such as participatory action research,[13] community-based participatory research,[14] and action science,[15] which draws on critical ways of knowing for greater relevance and affinity to the social realities of the local communities. New driving forces and trends such as the open movement and open information infrastructures, along with the increasing "datafication" and "platformization" of society, have had and will continue to have significant impacts on the future of citizen science.[16]

The ensuing data revolution, with its new developments in ubiquitous and pervasive computing technologies, has invigorated the practice of citizen science. Contextualizing the discussion of citizen science in the datafication regime, whereby many aspects of life are transformed into digital data that have value,[17] brings to attention the capacity of citizen science as a data-intensive and data-driven field.

As a domain of data and a research method, citizen science can be investigated from its epistemological and ethical dimensions. So far, in scholarly work around citizen science, the focus has largely been on the benefits and challenges of citizen science,[6] the level of participation of citizens in citizen science (see Table 1), the contribution of citizen science to filling the data gaps,[18] the quality of scientific discoveries of citizen science to fulfill its promises on the science side,[19] the value and usability of citizen science data,[20] and the power dynamics between professional researchers (as experts and project facilitators) and participants.[21]

From the standpoint of critical data scholarship, the published work in citizen science currently focuses on the ethical considerations of citizen science (see, for example, Chesser et al. and Scheibner et al.).[22,23] However, as a data-driven field in which power asymmetries are inherent, there is a need to explore the wider structural issue within citizen science by investigating the broad sociotechnical assemblages embedded in the data and technology infrastructures that shape and are shaped by citizen science. In this regard, examining the attainment of the different elements of social justice, including participation, equality, equity, representation, and accountability, within the domain of citizen science is imperative.

Situating citizen science at the interplay between data-driven processes and social justice issues offers a way to fulfill the two exigencies and address the literature gap in citizen science. To this end, a "data justice" framework is used as an analytical lens in this work. The data justice concept builds from concerns regarding the adverse impacts of datafication on individuals and groups.[24,25] Framing citizen science from the data justice perspective provides a lens to identify the nature and diversity of social justice and the social injustice emanating from the data processes in citizen science to highlight the significance of the contexts of citizen science as a data-based practice. Utilizing this framework, this work addresses the literature gap in citizen science by affording a window into investigating the exploitative and exclusionary implications of data-driven systems of citizen science on citizens. In this regard, this work contributes to the scholarship of citizen science and critical data studies.

Critical data scholars have studied datafication from the viewpoint of data justice and pointed out several implied
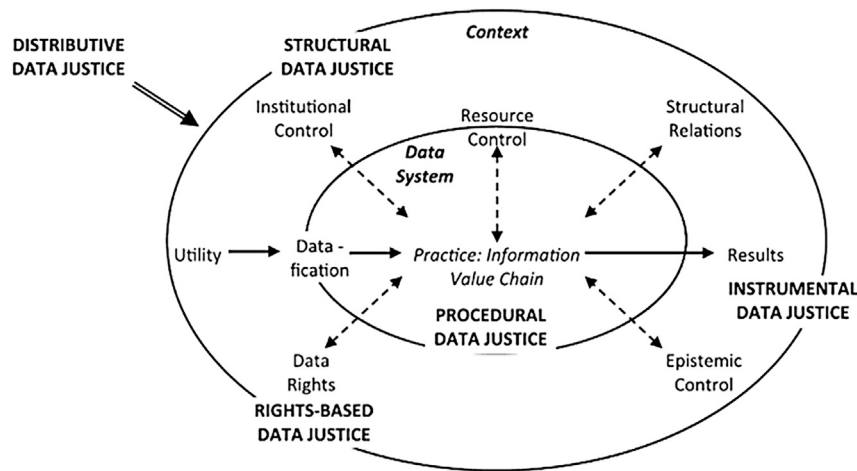
comprising the upstream steps of data capture, midstream steps of data processing into information, and downstream steps of using information for shaping decisions and actions, leading to the generation of value as the results of the information value chain (see Figure 2).

The first dimension, procedural data justice, assesses fairness in the system's handling of data within the information value chain. The second one, instrumental data justice, focuses on the outcomes of data use, and thus fairness in the results of data processes. The third dimension, rights-based data justice, relates to basic data rights, including the rights of data access, ownership, privacy, and representation. Structural data justice considers the social structures that shape and are shaped by the data systems and data-related outcomes. Finally, distributive data justice is an overarching lens that encompasses all other dimensions and relates to a broader concern of equity of data-related resources. Distributive data justice considers that for a fair distribution of data results and just data functioning to be achieved, there must be a fair distribution of data-related resources.

The rest of the article (1) briefly reviews the selected citizen science cases, (2) investigates the five dimensions of data justice in the data processes of the citizen science cases, and (3) discusses the relevance of data justice as an analytical framework to the field of citizen science.

Over the years, many citizen science projects have been undertaken across different fields and across many countries. Rather than providing a comprehensive overview of the different types of citizen science initiatives, this article explores four citizen science cases. The selected cases are across the domains of personal data research, weather observation, water quality monitoring, and forest governance, as follows:

social justice risks, including the emergence of new forms of exclusion and inequality for vulnerable groups (see, for example, Donovan and Masiero and Das)[26,27]; unfairness in "the way people are made visible, represented and treated as a result of the production of digital data";[25] and constraints associated with lack of legibility, agency, and negotiability in data systems.[28] Meanwhile, despite extensive research into the scientific outcomes of citizen science and the assessment of participants' engagement in citizen science, little has been done to understand how the data processes of citizen science operate and support the achievement of social justice for the participants of citizen science, in particular, and those who are affected by it, in general.

To do this, we adopt Heeks and Shekhar's[29] data justice framework, which provides a conceptual foundation to explore the structural drivers that shape and are shaped by the data processes and data-related outcomes, the implications of mass data processes on citizen science participants and those affected by it, and potential mitigation strategies for addressing such implications from the design and process perspectives.

Data justice is "the primary ethical standards by which data-related resources, processes, and structures are evaluated."[24] The conceptualization of data justice in the framework is particularly of relevance to the discussion in this article as it outlines various dimensions in which to analyze data-intensive and data-driven processes with respect to structural inequality and social injustice. Data justice situates data-based processes like citizen science in the broader complex sociotechnical systems consisting of several interrelated elements of discourse, material resources, the political economy, institutions, and social relations, or what is called the "data assemblage."[30]

Understanding data justice and injustice from the viewpoint of datafication in the development sector, Heeks and Shekhar[29] conceptualize data justice as comprising five dimensions in relation to data flows within the "information value chain"[31] and the results of the data system: procedural, instrumental, rights-based, structural, and distributive data justice (see Figure 1). The information value chain is a model that represents the processes of data valorization from data inputs into results,

- Launched in 2015, Open Humans (OH; www.openhumans.org) is a community-driven data platform that facilitates personal data aggregation across several data sources. It was initiated and is run by the non-profit Open Humans Foundation. OH aims to leverage personal data to help develop empirical knowledge and enable participant-centered data exploration. The platform hosts projects such as open-source diabetes tools and applications for managing and visualizing diabetes data (e.g., Nightscout, OpenAPS), as well as online activities analysis (e.g., Google Search History Analyzer). OH also facilitates the donation of data-processing tools.[32] In OH, lay participants can make data contributions to projects, conduct self-research, and participate in the governance of the platform.[33] As the aggregator of data and participants, OH is a mediator within the data ecosystem.
- Old Weather (OW; www.oldweather.org) is a weather data crowdsourcing project based in the United Kingdom that
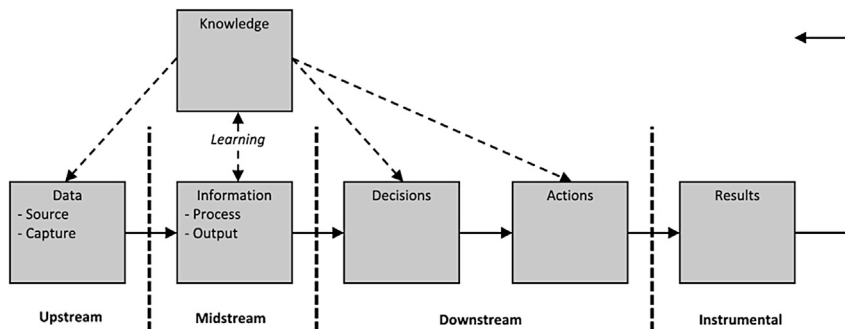
to capitalize on the commitment of the government of the Democratic Republic of Congo to the European Union's Forest Law Enforcement Government and Trade (EU FLEGT) Voluntary Partnership Agreement.[41]

These four cases were explored, through the analytical lens of the data justice framework, to unpack the social justice dynamics associated with the data systems embedded in these citizen science projects. For each of the cases, secondary sources were consulted and informed the analyses. As far as the data were available, they were used to support the observations across each of the data justice dimensions (i.e., procedural, instrumental, rights-based, structural, and distributive). The methods for case selection and analysis are further described in the Experimental procedures. See Table 2 for the summary of citizen science cases.

RESULTS

Across the four citizen science projects, citizen science lay participants assume the role of data contributors, while the administrators of the data platform and professional researchers assisting citizen scientists in the data processes are the data stewards who are responsible for managing the data processes in citizen science. The role of an intermediary who stimulates the flow of data between the data source and the data users and contributes to increasing the accessibility and utility of data[42] is assumed by project administrators in OH and OW; professional researchers, reporters, and activists in FWS; and non-governmental organizations in ECS. The beneficiaries, who benefit from the data processes, are lay participants across the four cases; professional researchers who use and reuse the citizen science data in OH; the UK Met Office and the organizations in the financial services industry that reuse OW weather data; policymakers, regulators, criminal justice processes, and intermediaries involved in FWS; and logging companies, law enforcers, researchers, and intermediaries in ECS.

Procedural data justice
The motivations and objectives of citizen science projects inform the design of the supporting data systems. The design of the systems has implications on the data-handling processes and the kind of functionalities (i.e., interactions between project stakeholders, activities, and the extent to which participants take part in the data processes) that are afforded by the data systems. Assessing procedural data justice thus requires understanding these objectives as well as the data processing that ensues in these systems.

OH aims to leverage the use of individual health data and online activities data to help grow empirical knowledge and enable

was launched by the Zooniverse citizen science consortium in 2010 and led by climate scientists at the UK Met Office. Citizen scientists (lay participants) were recruited to recover ship log accounts of epidemics, people falling overboard, and getting stuck in the ice, and the weather observations made by the crews of historic ships by transcribing digitized versions of ships' logbooks and uploading the transcriptions to OW's open data platform.[10] OW project administrators publish the transcription tasks on the website, from which citizen volunteers can choose to work on. These transcriptions are then used for modeling weather and climate projections and improving scientific knowledge of past environmental conditions.[10,34]

- The Flint Water Study (FWS; flintwaterstudy.org) is community-based participatory research that started as a response to the water quality crisis in Flint, Michigan, in 2014. Flint residents partnered with academics from Virginia Tech to produce credible evidence to support the residents' claims about public health and environmental threats resulting from poor water quality.[35] Trained by the Virginia Tech team and Flint community leaders, citizen volunteers performed systematic data collection of water samples across the city.[36] Sampling kits were also provided for citizens. The FWS found high levels of lead contamination in the water supply. Findings from the study have been used in legal and social advocacy.[37,38]

- Extreme Citizen Science—ExCiteS (geog.ucl.ac.uk/research/research-centres/excites) is a research group based at University College London that, among others, worked in the Congo Basin rainforest to support the Mbendjele Indigenous people in conducting community mapping of natural resources, recording illegal activities in the forest areas, and community engagement in forest governance starting in 2013.[39] The project (hereafter, ECS) started with the local community's request for the development of a tool for mapping key natural resources and recording illegal logging activities.[40] Through a participatory technology design process with the community members, the research group developed a resource-mapping device for the non-literate, unschooled community members. In partnership with local intermediaries, the research group trained community members to analyze data and use it in the advocacy effort. The project seeks

Table 2. Summary of citizen science cases

Factors	Open humans	Old weather	Flint water study	ExCiteS in DRC
Location	Global	UK	USA	DRC
Project aim	Leveraging personal data to help grow empirical knowledge and enable participant-centered data exploration	Crowdsourcing public participation in transcribing historical weather data to inform weather and climate modeling	Gathering evidence to support residents' claims about public health and environmental threats resulting from poor water quality	Mapping community resources and gathering evidence of illegal logging to support more equitable participation of local communities in the forest governance process
Domain	Personal informatics (e.g., behavioral data, health informatics)	Weather observation	Water quality monitoring	Forest governance
CS model category	Contributory to co-created (based on the project)	Contributory	Co-created	Co-created
Degree of participation	Nominal to representative (based on the project)	Nominal	Representative	Representative
Social justice dimension	• Representation: citizen participation in knowledge production • Representation: adherence to basic data rights	Representation: citizen participation in knowledge production	• Representation: citizen participation in knowledge production • Representation, redistribution, and recognition: using data to challenge structural injustices	• Representation: citizen participation in knowledge production • Representation, redistribution, and recognition: using data to challenge structural injustices • Representation, redistribution, and recognition: using data as a tool to challenge self-sense of powerlessness against powerful actors

participant-centered data exploration.[33] The design and governance of OH are based, among others, on considerations of practical problems often faced by citizens (e.g., how to merge data streams from various sources). Individuals determine which project in the OH platform they want to participate in and thus how their data are used.

The governance of data in OH is managed at multiple levels. Essentially, data contributors have full control over access to their data contribution. Data are uploaded by participants to their account, and access to participants' data is controlled through a granular consent procedure that asks participant permission to share potentially identifiable data with projects in the platform. All personal data stored in participants' accounts are accessible only to the participants themselves unless they choose to share them with specific research studies or make individual datasets publicly available.[33] It is thus observed that OH employs an explicit consent mechanism.

In addition, layperson participants as members of the OH community can participate in the review and approval process of the new projects that want to be shared on the OH site.[33]

They also get to elect some of the members of the Open Humans Foundation board of directors, enabling them to participate in the broader governance of the platform.[33,43] This configuration contributes toward enabling accountability and transparency in the use of data in such a way that participants' data contribution is not used beyond their intention, such as health profiling based on stereotypes for governance and commercial purposes and loss of privacy.

Using the functionality available in the platform, participants can explore and analyze their own data. This affordance extends participants' control over the data processes and data-related outcomes, as they also handle data analysis, interpretation, and the subsequent data processes. The platform also provides the opportunity for project administrators to request the inclusion of participants' data for reuse in the new research.[33] Therefore, any data use and reuse for value-adding purposes is determined by participants.

OW aims to facilitate crowdsourced data transcription. The project has enabled the mobilization of large amounts of data locked in analog records as digital information and uses them

in previously impossible ways.[44] These transcriptions contribute to weather and climate model projections and improve scientific knowledge of past environmental conditions. This information is used by researchers around the world.[10]

The design of the OW's data system allows lay participants to act only as data transcribers and data editors, while the rest of the data processes are handled by the project administrators and professional researchers/scientists. Meanwhile, the identification of transcription tasks to complete, data processing, and data use and reuse are determined by the institutions administering the project, which include the Met Office, the National Archives, the National Maritime Museum, and the National Oceanic and Atmospheric Administration. The data processes in OW are thus somewhat "extractive"[29,45] in the sense that data gathering by citizens benefits the aforementioned agencies, who have the resources to use the information, while citizens themselves do not directly benefit from their labor. This is a common feature of contributory citizen science projects that need to collect large datasets to gain a deeper understanding of natural phenomena and rely on microtasking and light engagement.[46,47]

Both OH and OW employ data platforms to which participants upload their data contribution. Whereas OH allows participants' deep involvement in the data-handling processes, the architecture of OW's data platform strictly limits participants' involvement in the said processes. There are greater procedural benefits and process gains for OH's participants than for those of OW.

The objective of the FWS is to conduct systematic investigations to the quality of water in the city of Flint and its evolving impacts on residents.[35] Residents were given instructions to conduct voluntary sampling of their water and submit these samples to Virginia Tech for analysis.[36] In that sense, the data were captured by citizen science participants, but the epistemic conversion of data to knowledge and applied outcomes for data-driven decisions and actions was handled by researchers. From there, citizen science participants were involved in the subsequent data processes alongside intermediaries, including the scientists, activists, and civil society organizations (CSOs).[48] Results of the data analysis have been made available online.

ECS in the Congo Basin rainforest aimed to equip the Indigenous communities with data collection devices that could support them in mapping their territories and key natural resources and recording illegal logging activities. Using the devices, they could collect the evidence that they needed for effectively participating in the collaborative forest-monitoring and governance processes alongside CSOs.[39,40]

The design of the data collection device took into account practical challenges specific to the Congo Basin rainforest community, including adverse conditions of the African rainforest, the lack of power facilities, and security risk in the form of participants being caught by eco-guards when conducting data collection.[9] Community members were involved in the midstream data value chain of analyzing collected data into usable information and the downstream process of seeking redress for violation of several aspects of forestry laws and their Indigenous rights with the help of CSOs.[39]

In both the FWS and the ECS, lay participants took part in all data processes, from shaping the research design and the upstream steps to downstream steps of the information value chain. However, compared with intermediaries, they were less involved in the midstream and downstream steps, and thus generated fewer procedural benefits and process gains in terms of knowledge of the decision-making processes (i.e., how data should be processed and presented for effective advocacy and how to engage with local administrations) and network-building. Through their participation in citizen science, participants may gain interpretive value in the forms of a sense of contentment for having contributed to the projects, a greater sense of legitimacy and authority as citizens, and a sense of belonging to the cause that is advocated through the projects.[36,40]

Instrumental data justice

Instrumental data justice focuses on the downstream processes of the information value chain—in the outcomes of data being used. It deals with the questions about who benefits from the use of data and, particularly, data injustices resulting from the data being used.

In the case of OH, lay participants can gain instrumental value by conducting self-research. They can work with their data and utilize the platform's affordances for exploring and analyzing their personal data, as well as for discovering interesting patterns across various data sources. Self-research allows them to gain applied insights about their health condition, lifestyle, and online activities in comparison with other people, and to conduct relevant actions upon such an understanding.[33] Participants who are more interested in exercising their goodwill toward advancing science through their data donation to studies hosted in the platform may gain pro-social value. OH prompts instrumental value for research projects hosted on the platform using project-specific data and reuse of public data.

OW can benefit participants by affording them the opportunity to gain knowledge about the weather conditions in the past as they participate in the transcribing and editing tasks. However, data use provides more instrumental value for the UK Met Office and, subsequently, those who reuse data, including the weather derivatives market, businesses, academics, and policy makers. Studies on OW note that participants' desire for a variety of sociocultural values motivates their engagement with the project.[10,49] These include a desire for connecting with people who share the same interests, a sense of responsibility to contribute to the community, and a desire for completeness and accuracy in their data contribution. Fulfilling the said desires could result in individual gains.

The UK Met Office, as one of the OW beneficiaries, values immediacy in the pace at which data are cleaned and homogenized more than accuracy and is more interested in the "attractive" weather modeling developed from the data contribution rather than the data work itself.[49] OW generates values for the UK Met Office and commercial users that reuse data outputs by giving them access to complete and accurate transcription that allows them to model weather and climate projections, as well as participants' sustained contribution.[44] Injustice then arises from using OW data that are produced through participants' free labor in ways that serve the political and economic interests of the UK Met Office and businesses, and simultaneously disregarding participants' concerns around the completeness, consistency, and correctness of OW data.

Further, participants reported that they were struggling to maintain motivation when the project incorporated new transcription tasks that participants were not familiar with and appeared to be intellectually and emotionally difficult to them.[10] This instance demonstrates how a citizen science project can keep on benefitting other actors and the research project through uninterrupted data collection, on one hand, while being less benefitting to participants on the other hand.

In FWS, data provided instrumental value for citizens in five ways. First, they were fed into citizens' decision-making to inform their advocacy strategy and legal redress.[36] Second, data offered empirical support for the Flint citizens to communicate on more equitable terms with policy makers and regulators who considered data-based knowledge a more accurate and trustworthy source of knowledge compared with citizens' narratives.[37] The FWS data negated misleading official data, which the state and local officials used to defend their policy, and became scientific evidence to support Flint citizens' class action lawsuits.[36] Last, data helped citizens generate support from concerned members of the public, including activists and donor organizations.[38,48]

The FWS data validated citizens' claim against the state's crimes. The data eventually enabled the service of justice for citizens as demonstrated in the instances where the city of Flint received federal and state funding for pipe replacement, health, and education;[50] some public officials were charged with involuntary manslaughter;[51] and the US Supreme Court ruling allowed the Flint residents to sue government officials over the city's water crisis.[52] However, a form of structural injustice persists as the Michigan emergency management law that gave rise to the crisis remains in place.

ECS aimed to support the local community to advocate for their participation in the governance of natural resources. Through the development and use of technology for data collection and data visualization, the project enabled the Congolese rainforest community members to map their resources and gather and report feedback on illegal logging activities to the forestry sector watchdog, and support themselves in seeking redress for violations of their rights to lands, territories, and other natural resources as recognized by law.[39,40] The activity thus has made visible the living environment of the local community to external actors. This "external visibility" brought some instrumental values: the more representative map of the Indigenous lands and resources and the communities' better understanding of their territories.[39]

The map generated from the project has helped to inform logging companies that attempted to comply with the EU FLEGT requirement specifying that only legally harvested timber can be exported to the EU.[53] To avoid violating the Indigenous Peoples Law, companies sought to respect the rights and resources of Indigenous and local forest people in their logging activities by not interfering with resources that were critical to the community's livelihood.[41]

The project also brought instrumental value to CSOs whose work focused on monitoring illegal logging and helping local communities seek redress mechanisms through the Independent Monitoring-Forest Law Enforcement and Governance approach.[54] The empowerment of CSOs in their forest monitoring work transformed their profile as an active contributor to improve forest governance.

Across the four cases under review, there are more potentials for lay participants in OH, FWS, and ECS to gain instrumental value from the outcomes of citizen science projects compared with OW participants. The use of data in OW and its outcomes bring more instrumental value for external agencies than to the lay participants. In FWS and ECS, data could act as a tool to alleviate power inequalities between citizens and government and law enforcement by making visible violations of citizens' rights and injustices and hold the government and untrustworthy law enforcement accountable. For providing new pathways of cooperation and mutual learning space for the actors involved in the project, OH is likely to equally benefit lay participants, the project administrators, and their research.

Data generated from the four projects and the outcomes empower not only lay participants, but also intermediaries, who use the data in their advocacy, engagement with policy makers, and interventions. Professional researchers may obtain symbolic resources, such as prestige and reputation, from their participation in the projects and the publications resulting from the use of citizen science data.[45]

Rights-based data justice

Rights-based data justice focuses on the commitment to basic data rights, including the rights of data access, ownership, privacy, and representation. Whereas, within the domain of data justice, the four rights are framed in relation to the use of personal and privacy-sensitive data, the data that are collected and used in citizen science are not limited to this type of data. Consequently, the discussion on data rights can be extended to cover any form of citizens' data contribution whose production and collection rely on citizens' labor.

Given the architecture of OH's data platform and its privacy-preserving protocols, participants have control over data access and ownership, as well as control over their external visibility in the datasets. Further, participants have the right to have their previously recorded data revoked and deleted from projects and the platform, or what is known as the "right to be forgotten."[55]

The use of metadata as the by-products of participants' interactions with the platform and projects needs to be done with participants' consent; further, they are made aware of how their data contribution will be used in the projects. Through its iterative approaches to data sharing, OH enables the processing of participants' personal data in accordance with their needs and interests,[33] thus mitigating the risks of genetic discrimination, loss of privacy, and reidentification in publicly shared data. The opt-in model implemented in OH requires participants to give a granular consent for data sharing and use for every project in the platform,[33] thus making each potential data use mediated by each research study. Further, the midstream data processes in OH are visible to the participants.

OW does not gather and use personal or privacy-sensitive data. Therefore, the issue of data privacy and representation within OW does not have the same weight as it does in citizen science projects that use such data. Participants have access to their data contribution and this individual contribution is recorded and identified in the transcriptions. This gives participants a sense of ownership over their data contribution and the project.[56] However, OW participants are generally not aware

that the outputs of their contribution are used not only for meteorological, climatological, and geophysical purposes, but also by the weather derivatives industry.[44] In this regard, an informational asymmetry may exist between the institutions administering OW and the participants. This lack of transparency in the data processes following data upload creates injustice in terms of the right of ownership of data contribution.

In FWS, participants collected samples and sent them to the researchers for analysis. They would then have access to the outputs of data processes that were uploaded to the FWS website. The FWS data rendered the "worst case" homes in Flint that were served by lead pipes more visible compared with the representation in the official data,[36] thus giving a more accurate representation of the community in data and making the magnitude of the crisis more apparent. In exchange for representation in data, citizens' ownership of data that captured some aspects of their living conditions was ceded to researchers who conducted data analysis and who, to some extent, had authority on the narrative resulting from the analysis results.

In 2018, the Flint residents filed a complaint (flintcomplaints.com) against the lead researcher of the FWS, Professor Marc Edwards of Virginia Tech, for misrepresenting the Flint residents in Edwards's complaint against the Flint Area Community Health and Environment Partnership, violating the residents' right of self-determination, defamation of the residents, and false claim about the safety of Flint's residential water. That served as an instance of rights-based injustice whereby the right of the Flint residents to be fairly represented in data-related outcomes was violated during the midstream steps of the value chain. In this regard, while fairness in data representation in the upstream processes of FWS could be served, the same could not be guaranteed in the subsequent processes.

The data that were used in ECS were spatial data and participant observation data that were gathered through the observation conducted by community members.[9,39] In ECS, community members conducted the data collection themselves and were assisted by intermediaries in the midstream data processes of making sense of the visualization of their observations and in encouraging data-utilizing activities by governments. These intermediaries had more skills and expertise in translating data into relevant information for decision-makers. Community members had access to the data that they contributed; however, the downstream steps of the information value chain of the data-based decision-making process were largely invisible to them.[39] Data made the living conditions of the local communities and the environmental threats that they faced more visible to decision-makers. This visibility, however, came with a risk of the mapping approach being abused by the elites to give their executive decisions some appearance of popular legitimacy.

Related to visibility in data, while there was a risk of local communities being spotted or caught while making observations, the data collection instrument used in ECS mitigated the concern in the design process of the instrument. The instrument enabled observation contributors to provide an estimate of where the illegal activity occurred without being in the exact location of the actual occurrence.[9]

Whereas OH adheres to the four basic data rights, the midstream and downstream steps of data in OW produce injustice in terms of the right of data ownership. In FWS and ECS,

while lay participants had access to the upstream data processes, the downstream data processes were invisible to them. Therefore, unless data intermediaries in the FWS and ECS took up specific actions to counter this invisibility, participants were unable to either participate or monitor data-related outcomes, i.e., actions and decisions made about them. In both cases, representation in data implied ownership of data and data-related outcomes was shifted from the local communities. It also carried the risk of misuse of data for misrepresenting the community (as in the case of FWS) and for primarily serving the interest of powerful institutions (as in the case of ECS).

Structural data justice

Structural data justice focuses on how structures shape data systems and how data systems shape structure. The social structures that shape and are shaped by data systems include structural relations, epistemics, utility, and institutions.

Structures shaping data systems

Social structures are partly responsible for enabling and constraining data uses, processes, distributions, and rights.[57] In citizen science, relations of power are mostly understood in terms of the relationship between lay participants and professional researchers/scientists as experts and project administrators. However, there are a lot more actors with different needs and interests involved in a citizen science project.

Structural relations or relations of power can be broadly understood as the relations between those who use data outputs for decisions and actions (e.g., local and national governments, law enforcers) and the lay citizen science participants for contributing to the achievement of outcomes for sociopolitical, economic, and ecological systems, such as policy influencing and community capacity building for decision-making.[58]

Informed by data rights discourse, OH adopts privacy-preserving protocols, resulting in data processes that place control over data and most of its outcomes in the hands of participants. This practice challenges research with biomedical data, which has traditionally failed to give patients power over how their data can be used and carried risks of loss of privacy and reidentification in publicly shared data,[59] as well as participatory citizen science health research that focused on crowdsourcing data from participants to support scientists without giving them access to the data they contribute. Further, the actual usage of data by researchers and participants serves the agenda of justice as it provides utility for both stakeholders, without the abuse of power by researchers and platform administrators over participants' data contribution and the overall data system. There are more spaces for lay participants' contribution in the data governance of OH compared with the other cases under review in this study.

For OW, the use of weather observations is shaped by the interest and agenda of the UK Met Office.[44] The actual usage of data by the Met Office results in better climate and weather modeling, and thus indirectly benefits participants, whose motivation to contribute to science and the larger social contexts is being fulfilled. However, as found in the study of Bates et al.,[34,44] the data user appropriates participants' free labor and the knowledge generated from the project to serve its

interest to generate profit from selling data and other value-added information products to commercial users. OW thus maintains the power differentials between participants and its hosting agency that dictates the data processes.

In terms of epistemics, the discourses of Big Data and neoliberal New Public Management, which has been adopted by the UK public sector, inform the transformation of OW's data into data-related outcomes that benefit the UK Met Office.[49] As OW data contribution becomes part of the Met Office Big Data infrastructure, the use and reuse of data occur within a technocratic space that values utility, functionality, and economic reasoning more than participatory knowledge production. In a similar vein, the neoliberal New Public Management trend denoted by the restructuring of government operations along the market lines informs the treatment of OW data outputs as commercial assets to be monetized instead of as a commons. Although the UK Met Office, which hosts the citizen science project, is publicly funded and citizen science provides free labor for the production of data, access to the outcomes of OW data processed in the Met Office is free only for academics. This in turn hampers the distribution of knowledge.

In FWS, the government's monopoly of legitimate citizens' complaints about water safety and the erosion of local democratic accountability to citizens shaped data use. There is also evidence of environmental racism in the dismissal of citizens' complaints about water quality by officials responsible for emergency management.[60,61] Despite the presentation of evidence from FWS to the city and state officials, data were not used to inform decision-making, as they kept on refusing to acknowledge the credibility of citizens' evidence-backed claims, thereby delaying the undertaking of necessary actions and the declaration of a state of emergency to reduce the impact of the crisis.[48]

In terms of epistemics, the strength of neoliberal governance discourse shaped the data processes in FWS. This, for example, was exemplified in the city and local officials' characterization of health problems as individualized and as requiring the response of personal medical care rather than investments in public infrastructure.[48,62] It was further exemplified in the dismissal of citizens' widespread concerns over the poor water quality and its impacts on the public health as anecdotal claims and therefore invalid.[37]

The structural relations between professional researchers and Flint residents also shaped the data processes. The residents' complaint against the lead researcher previously discussed in the rights-based data justice section showcased the power imbalance between scientists and the disenfranchised communities whereby scientists' resources (i.e., knowledge, expertise, outreach) vis-à-vis those of Flint residents shaped the use of information for promoting data-utilizing activities by the government, resulting in the feeling of injustice on the citizens' part.

For ECS, data collection and use by Indigenous community members and intermediaries for better engaging in the forest and national resource governance were informed by an emerging Indigenous governance discourse and conservation paradigm. Institutional forces also shaped the data systems. Forest governance monitoring agencies, civil society networks, and compliant logging companies provided an enabling environment for the use of data to inform just forest governance.[9] However, this was constrained by inadequate law enforcement,

corrupt local officials, and logging companies seeking to harvest more trees than they were legally allowed.[39] In ECS, the structural relationship between the local communities and the researchers did not have a detrimental impact on data use, as the researchers were more involved in evaluating the usability of the data collection devices and not in the midstream and downstream data processes.[40]

It can be observed that utility, epistemics, structural relations, and institutional forces have different impacts on the data processes in citizen science projects. Where data have utility for users, there is greater acceptance to data-utilizing activities. In terms of epistemics, discourses rooting from the field of human rights and social justice, such as data rights in OH and Indigenous governance in ECS, support the use of data-related outcomes for attaining social justice agendas, while neoliberal discourse surrounding data processes in OW and FWS constrains that. Institutions and unequal relations between actors result in the form of exploitation in OW and resistance in using data-related outcomes in FWS and ECS.

Data systems shaping structures

Data processes in OH challenge dominant scientific practice with biomedical data by introducing an alternative way to generate quality health data and ethically use data by centering participants' consent across the steps of the information value change. This reflects a paradigm shift related to the relations between scientists, as data users, and research participants, as data contributors, moving in the direction of participants being the subject and center of research. Increasingly, medical researchers and organizations have started to adopt this privacy-based data system and use the data generated from the OH for research.

In terms of utility, OW has become an exemplar for a successful interdisciplinary weather monitoring project by enlisting the help of citizens for digital transcription of weather records. This reflects a shift in the recognition of the potential of citizen science for contributing to the domain of weather and climate science.

FWS and ECS mainly aimed to change the power relations between those in power and citizens. In FWS, was between the city and state officials and the citizens, while in the ECS case, it was between the forestry government, the actors conducting illegal logging and commercial poaching, and the citizens.

In FWS, where official data used in the decision-making proved to be unreliable and misleading, the data produced through a collaborative effort from citizens and professional researchers brought with it an interpretive value for devaluing official data and discrediting the decisions and actions made based upon it. Findings from the research validated citizens' claims and modified the perceived interests of decision-makers, resulting in the water supply change in October 2015[48] and Michigan officials facing criminal charges in 2017.[36] The widespread acceptance of the research findings also translated into public support for the advocacy.[38] This shift occurred because the quantitative language in the test results substantiated the qualitative local concern. Aside from external actors, the residents also gained better knowledge about the severity of the water crisis.[63] Despite evidence of institutional change resulting from FWS, underlying structural issues that gave rise to the crisis remain in place.[60]

For ECS, the mapping of tribal lands, community resources, and documentation of illegal logging and commercial poaching practices provided the community members of the Congo Basin rainforest with evidence to participate in the collaborative monitoring and management of local resources.[39] The community-owned maps became a new language through which communities could make their concerns known to the logging companies and local governments regardless of their literacy and educational background.[39] As a result of collaborative monitoring, participatory mapping with Indigenous peoples has become a standard practice for companies obtaining certification in Congo Basin and companies are required to respect the rights and resources of Indigenous and local forest people.[64]

Although a substantial improvement in the forest governance had not yet materialized, the government's growing commitment to tackle illegal logging was evident from an improved legality assurance system and the facilitation of CSOs in forest monitoring.[65] Further, there has been a growing private sector engagement in efforts to promote legal timber trade.[54]

Across the four cases, datafication through the means of citizen science has resulted in epistemic change whereby citizen-generated data are regarded as a valuable source of information and knowledge with merit for scientific research and decision-making.

DISCUSSION

Distributive data justice encompasses all the other dimensions of data justice and relates to the fair distribution of data-related resources and the results of data systems. As such, this dimension provides an overarching analytical lens to observe data injustices.

In procedural terms, except for OW, in which participants are involved only in the upstream steps of the information value chain, participants generally gain procedural benefits through their closer involvement with the project across data processes. However, compared with lay participants, researchers and project administrators who shape the data processes from the very beginning, albeit to different degrees, gain more procedural benefits. Across the four cases, participants may gain an epistemic value in terms of greater awareness of issues addressed in the projects and an interpretive value in the form of satisfaction from contributing to the projects. Although there were process gains for FWS and ECS participants, who acquired new data collection skills, intermediaries collaborating with participants in the midstream and downstream data steps may obtain more process gains, such as improved skills in using data for advocacy and contacts with decision-makers.

In terms of instrumental justice, participants' gains of instrumental value differ across projects. Participants in OH and FWS may gain more instrumental value from data compared with those in OW and ECS, which provide more instrumental value to actors other than participants. The privatization of data outcomes by the UK Met Office produced from the free labor of lay participants creates the conditions for distributive injustice, as the benefits accrue to the UK Met Office, which uses data for commercial purposes. In OH, lay participants and professional researchers may gain equal value from data

and data processes, as OH allows for a more balanced power relation between both stakeholders compared with the other three projects. Meanwhile, professional researchers and intermediaries collaborating with participants may gain all the aforesaid values throughout their involvement in facilitating and guiding the whole data processes.

Within the dimension of rights-based data justice, OH adheres to the four basic data rights. Although FWS and ECS facilitated participants' representation in data, this visibility carried risks related to data ownership, misrepresentation in data-related outcomes by scientists who controlled the midstream steps and intermediaries who were more active in encouraging data-utilizing activities, and misuse of data against their intended purposes by powerful actors. Further, participants in both projects did not have access to the downstream data processes in which decisions and actions were made about them. In terms of OW, the hosting agency controls access to and ownership of data contributions once they are uploaded.

Structurally, except for OH, it is evident that the structural inequalities of power-interest surrounding the data systems shape the data processes and outcomes. This is particularly the case for OW and FWS, whereby institutional forces, discourses, the utility of data to users, and structural relations have an impact on the information value chain, resulting in data processes and data use that align more with the interest and agendas of powerful actors than with those of participants in the citizen science projects. In ECS, corrupt officials, inadequate law enforcement, and non-compliant logging companies constrained the use of data to support justice-serving agendas.

Data have limited impacts on the structure. Data support the empowerment of citizen science participants vis-à-vis professional researchers and traditional scientific practices across the four cases, thus demonstrating a degree of epistemic impact. The role of data in supporting advocacy for justice and equity to communities fraught by unjust governance systems is, however, inadequate, particularly in the case of ECS. At best, the substantial impact of data availability is on building local communities' profiles as partners with merit to participate in the decision-making process. However, data have not had a significant impact on shaping the broader structural issues.

Across the dimensions of data justice, citizen science projects under review facilitate the attainment of social justice agendas in terms of involving citizens in scientific practice through which they can contribute to the production of knowledge, thus challenging the dominant cultures of knowledge production and, in the case of FWS and ECS, countering the invisibility of marginalized populations' living conditions through data. However, the utility of citizen science data in addressing structural injustice and tackling unequal distribution of power and data-related outcomes among actors within the data processes is limited.

In this study, we undertook an investigation of the data system and data processes of citizen science from the lens of data justice. The employment of data justice as an analytical lens has enabled an investigation into the multiple ways through which citizen science challenges and sustains injustices across the information value chain. As observed from the analysis, the application of the framework to the field of

citizen science should carefully consider the specificity of citizen science as a data practice to mitigate the risks of losing the nuance between the different forms of citizen science and misconstruing the affordances of citizen science to support the interests of justice and equality for the participants and those affected by it.

OW is an example of citizen science in which fairness concerning data means "creating the greatest utility for the greatest number,"[24] while the FWS and ECS are examples of citizen science in which fairness concerning data implies creating utility for those who were marginalized or most vulnerable in the society. Fairness in OH, on the other hand, could imply creating utility either for the most data projects hosted in the platform or for an underserved community of patients who use the platform to conduct self-research. The opportunities for citizen science to facilitate data processes that support the quest for justice and equality, therefore, need to be understood with respect to the motivation, design, and objectives of the citizen science projects.

FWS and ECS used quality of life indicators data and data that emerged from community mapping of the living environment with the intended outcomes for and impacts on the broader sociopolitical, economic, and ecological systems to challenge injustices. Both projects demonstrated the attainment of social justice agendas using citizen science data to redress power imbalances and call for institutional change. In the case of OW and OH, the social justice agenda of equity to participate in scientific research is served, provided that the public can equally contribute to scientific knowledge and the lay participant-professional researcher boundaries are overcome.

In our analysis, we found that more citizen participation in data initiatives and more representation in data do not necessarily change citizens' access to justice. Compared with citizens as participants of citizen science, other beneficiaries, including professional researchers and elected government officials, benefit more from data processes and outcomes.

The discussion in this article is based on the review of relevant evidence presented in secondary sources of the selected citizen science cases and the online platform of the citizen science projects. Accordingly, there might be aspects of the citizen science data processes that are not captured in these sources. Undertaking a systematic review of citizen science projects according to the full scope of all the dimensions of data justice using primary sources is thus our future research agenda.

EXPERIMENTAL PROCEDURES

Resource availability
Lead contact
Debora Irene Christine is the lead contact for this study and can be contacted by email at debora@unu.edu.
Data and code availability
This study did not generate any new datasets.
Materials availability
This study did not generate any materials.

Method details
Method 1: Scoping review
Using purposive sampling, the cases were selected with consideration of a mix of different levels of engagement and participation from professional scientists and citizen science lay participants, intended outcomes, dimensions of social justice, and the types of data collected and used, as well as the data and technology infrastructures employed. The selection was also informed by the availability of supporting literature and secondary data. A scoping review was conducted to identify these secondary sources. The extent to which the analysis of the data justice dimensions was conducted on the selected cases was thus contingent on relevant evidence presented in these sources and in the online platform of the citizen science projects.

Method 2: Classification of citizen science cases
Many attempts to draw typologies of citizen science have been made. Shirk et al.[58] build on the role of the public in the scientific process to suggest five project models for public participation in science (see Table 1). The five models, from contractual to the collegial contribution, describe an increasing gradient of involvement of the public. The model considers that the degree to which citizens participate and the quality of that participation are related to the achieved outcomes, i.e., outcomes for research, individual participants, or socioecological systems.[58]

Whereas the degree of participation is defined as the extent of citizens' participation in the scientific research process, the quality of participation concerns the extent to which the project's goals and activities "align with, respond to, and are relevant to the needs and interests of participants."[58]

The degrees of participation in this model generally align with other typologies of participation, including White's,[66] where engagement ranges from nominal participation of citizens (e.g., in data collection) to transformative participation, where citizens drive the project agenda and manage the projects.

To assess the quality of participation, the following elements are considered: inputs (i.e., the interests of the actors participating in the project), activities (i.e., the tasks relevant to developing project infrastructure, project management, and communication with collaborators), outputs (i.e., the initial results of activities), outcomes (i.e., the tangible results deriving from specific outputs), and impacts (i.e., long-term and sustained changes of the well-being of humankind or the ecosystem).

Based on the typological frameworks in Table 1, the projects hosted in the OH platform could fall in the contributory/nominal, collaborative/instrumental, or co-created/representative category. OW falls under the contributory/nominal category, while ECS and FWS fall under the co-created/representative category. Further, the cases are also associated with varying degrees of participation and concerned with different elements of social justice, including participation, representation, and democracy (see Table 2).

To guide the effort to understand the different social justice dimensions of selected citizen science cases, Fraser's[67] theorization of social justice provides a useful framework. Fraser argues that social justice presupposes parity of participation according to normative "social arrangements that permit all to participate as peers in social life"[68] and that it requires three conditions: redistribution, recognition, and representation. Redistribution refers to equal distribution of economic resources. Recognition refers to equal social standing according to cultural markers and social differentiation. Misrecognition occurs when individuals or groups are unable to participate in society due to structures of subordination based on cultural values and social stratification. Representation concerns who counts, who belongs, and who is included in the politicospatial domain of the society. Disparity in representation, on the other hand, leads to unequal democratic and procedural access to participation.

The participatory feature of citizen science, as well as the utility of citizen science data for counteracting the injustice of external invisibility, is associated with the three conditions. Citizen science facilitates the co-production of knowledge through scientific research, thus allowing people's representation in data and science. The inclusivity in citizen science allows for the needs and interests of public participants to be elevated in the otherwise exclusive domain of scientific efforts where the participation and concerns of lay participants have been historically marginalized.

Citizen science data can be used to counter the otherwise invisible realities of peoples' lives and surroundings due to the lack of data on certain population groups or statistical invisibility. This invisibility is often a major factor in political exclusion and marginalization. In this regard, the outcomes of citizen science can support the attainment of redistribution and/or recognition. Likewise, misrepresentation in data could result in misrecognition and/or maldistribution. By outlining the different layers of social justice in the context of datafication, we could identify the social justice dimensionality in the selected citizen science projects.

REFERENCES

1. Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V., and Shirk, J. (2009). Citizen science: a developing tool for expanding science knowledge and scientific literacy. BioScience 59, 977–984.

2. Goodchild, M.F. (2007). Citizens as sensors: the world of volunteered geography. GeoJournal 69, 211–221.

3. United Nations. Universal declaration of human rights. https://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf.

4. UNECE (1999). Convention on access to information, public participation in decision-making and access to justice in environmental matters. Int. Leg. Mater. 38, 517–533.

5. McKinley, D.C., Miller-Rushing, A.J., Ballard, H.L., Bonney, R., Brown, H., Cook-Patton, S.C., Evans, D.M., French, R.A., Parrish, J.K., Phillips, T.B., et al. (2017). Citizen science can improve conservation science, natural resource management, and environmental protection. Biol. Conserv. 208, 15–28.

6. Conrad, C.C., and Hilchey, K.G. (2011). A review of citizen science and community-based environmental monitoring: issues and opportunities. Environ. Monit. Assess. 176, 273–291.

7. Zevin, M., Coughlin, S., Bahaadini, S., Besler, E., Rohani, N., Allen, S., Cabero, M., Crowston, K., Katsaggelos, A.K., Larson, S.L., et al. (2017). Gravity spy: integrating advanced LIGO detector characterization, machine learning, and citizen science. Class. Quan. Gravity 34, 064003.

8. Cooper, C.B., Dickinson, J., Phillips, T., and Bonney, R. (2007). Citizen science as a tool for conservation in residential ecosystems. Ecol. Soc. 12, art11.

9. Vitos, M., Lewis, J., Stevens, M., and Haklay, M. (2013). Making local knowledge matter: supporting non-literate people to monitor poaching in Congo. In Proceedings of the 3rd ACM Symposium on Computing for Development (ACM DEV '13; Association for Computing Machinery), pp. 1–10.

10. Lin, Y.-W., Bates, J., and Goodale, P. (2016). Co-observing the weather, Co-predicting the climate: human factors in building infrastructures for crowdsourced data. Sci. Technol. Stud. 29, 10–27.

11. Gieryn, T.F. (1983). Boundary-work and the demarcation of science from non-science: strains and interests in professional ideologies of scientists. Am. Sociol. Rev. 48, 781–795.

12. Haraway, D. (1988). Situated knowledges: the science question in feminism and the privilege of partial perspective. Fem. Stud. 14, 575–599.

13. Baum, F., MacDougall, C., and Smith, D. (2006). Participatory action research. J. Epidemiol. Commun. Health 60, 854–857.

14. Wallerstein, N.B., and Duran, B. (2006). Using community-based participatory research to address health disparities. Health Promot. Pract. 7, 312–323.

15. Argyris, C., and Schön, D. (1991). Participatory action research and action science compared: a commentary. In Participatory Action Research, W.F. Whyte, ed. (SAGE Publications, Inc.), pp. 85–96.

16. Schade, S., Granell, C., Vancauwenberghe, G., Keßler, C., Vandenbroucke, D., Masser, I., and Gould, M. (2020). Geospatial information infrastructures. In Manual of Digital Earth, H. Guo, M.F. Goodchild, and A. Annoni, eds. (Springer), pp. 161–190.

17. Mayer-Schönberger, V., and Cukier, K. (2013). Big Data : A Revolution that Will Transform How We Live, Work, and Think (Houghton Mifflin Harcourt).

18. Chandler, M., See, L., Copas, K., Bonde, A.M.Z., López, B.C., Danielsen, F., Legind, J.K., Masinde, S., Miller-Rushing, A.J., Newman, G., et al. (2017). Contribution of citizen science towards international biodiversity monitoring. Biol. Conserv. 213, 280–294.

19. Reason, P., and Bradbury, H. (2008). Editorial. Act. Res. 6, 5–6.

20. Fraisl, D., Campbell, J., See, L., Wehn, U., Wardlaw, J., Gold, M., et al. (2020). Mapping citizen science contributions to the UN sustainable development goals. Sustain. Sci. 15, https://doi.org/10.1007/s11625-020-00833-7.

21. Keune, J.D. (2019). Considering power relations in citizen science. Am. J. Bioeth. 19, 48–49.

22. Chesser, S., Porter, M.M., and Tuckett, A.G. (2020). Cultivating citizen science for all: ethical considerations for research projects involving diverse and marginalized populations. Int. J. Soc. Res. Methodol. 23, 497–508.

23. Scheibner, J., Jobin, A., and Vayena, E. (2021). Ethical issues with using internet of things devices in citizen science research: a scoping review. Front. Environ. Sci. 9, art629649.

24. Heeks, R., and Renken, J. (2018). Data justice for development: what would it mean? Inf. Dev. 34, 90–102.

25. Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. Big Data Soc. 4. 205395171773633. https://doi.org/10.1177/2053951717736335.

26. Donovan, K.P. (2012). Seeing like a slum: towards open, deliberative development. Georget. J. Int. Aff. 13, 97–104.

27. Masiero, S., and Das, S. (2019). Datafying anti-poverty programmes: implications for data justice. Inf. Commun. Soc. 22, 916–933.

28. Mortier, R., Haddadi, H., Henderson, T., McAuley, D., and Crowcroft, J. (2014). Human-data interaction: the human face of the data-driven society. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2508051.

29. Heeks, R., and Shekhar, S. (2019). Datafication, development and marginalised urban communities: an applied data justice framework. Inf. Commun. Soc. 22, 992–1011.

30. Kitchin, R., and Lauriault, T.P. (2018). Toward critical data studies: charting and unpacking data assemblages and their work. In Thinking Big Data in Geography: New Regimes, New Research, J. Thatcher, J. Eckert, and A. Shears, eds. (University of Nebraska Press), pp. 3–20.

31. Heeks, R.A. (2017). Structural model and manifesto for data justice for international development. SSRN Electron. J. Develop. Inform. https://doi.org/10.2139/ssrn.3431729.

32. Bietz, M., Patrick, K., and Bloss, C. (2019). Data donation as a model for citizen science health research. Citiz. Sci. Theor. Pract. 4, 6.

33. Tzovaras, B., Angrist, M., Arvai, K., Dulaney, M., Estrada-Galiñanes, V., Gunderson, B., Head, T., Lewis, D., Nov, O., Shaer, O., et al. (2019). Open humans: a platform for participant-centered research and personal data exploration. GigaScience 8, giz076.

34. Bates, J., Goodale, P., Lin, Y., and Andrews, P. (2019). Assembling an infrastructure for historic climate data recovery: data friction in practice. J. Doc. 75, 791–806.

35. Flint Water Study. The Virginia Tech Research Team. http://flintwaterstudy.org/about-page/about-us/.

36. Langkjær-Bain, R. (2017). The murky tale of Flint's deceptive water data. Significance *14*, 16–21.

37. Gaber, N. (2019). Mobilizing health metrics for the human right to water in Flint and detroit, Michigan. Health Hum. Rights *21*, 179–189.

38. Jahng, M.R., and Lee, N. (2018). When scientists tweet for social changes: dialogic communication and collective mobilization strategies by Flint water study scientists on twitter. Sci. Commun. *40*, 89–108.

39. Stevens, M., Vitos, M., Altenbuchner, J., Conquest, G., Lewis, J., and Haklay, M. (2014). Taking participatory citizen science to extremes. IEEE Pervasive Comput. *13*, 20–29.

40. Vitos, M., Altenbuchner, J., Stevens, M., Conquest, G., Lewis, J., and Haklay, M. (2017). Supporting collaboration with non-literate forest communities in the Congo-basin. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (ACM), pp. 1576–1590.

41. Broekhoven, G., and Wit, M. (2014). Linking FLEGT and REDD+ to Improve Forest Governance; ETFRN News Issue No. 55 (Tropenbos International).

42. van Schalkwyk, F., Willmers, M., and McNaughton, M. (2016). Viscous open data: the roles of intermediaries in an open data ecosystem. Inf. Technol. Dev. *22*, 68–83.

43. Kariotis, T., Ball, M.P., Tzovaras, B., Dennis, S., Sahama, T., Johnston, C., Almond, H., and Borda, A. (2020). Emerging health data platforms: from individual control to collective data governance. Data Policy *2*, e13.

44. Bates, J., Lin, Y.W., and Goodale, P. (2016). Data Journeys: Capturing the Socio-Material Constitution of Data Objects and Flows. Big Data Soc. *3*, 1–12.

45. Chen, S.-L. (2019). How empowering is citizen science? Access, credits, and governance for the crowd. East Asian Sci. Technol. Soc. *13*, 215–234.

46. Raddick, M.J., Bracey, G., Gay, P.L., Lintott, C.J., Murray, P., Schawinski, K., Szalay, A.S., and Vandenberg, J. (2010). Galaxy zoo: exploring the motivations of citizen science volunteers. Astron. Educ. Rev. *9*, 010103.

47. Reed, J.; Raddick, M.J.; Lardner, A.; Carney, K. An Exploratory Factor Analysis of Motivations for Participating in Zooniverse, a Collection of Virtual Citizen Science Projects. In 2013 46th Hawaii International Conference on System Sciences; 2013; pp 610–619.

48. Krings, A., Kornberg, D., and Lane, E. (2019). Organizing under austerity: how residents' concerns became the Flint water crisis. Crit. Sociol. *45*, 583–597.

49. Bates, J., Goodale, P., and Lin, Y. (2015). Data journeys as an approach for exploring the socio-cultural shaping of (Big) data: the case of climate science in the United Kingdom. In iConference 2015 Proceedings (Newport Beach) http://hdl.handle.net/2142/73429.

50. Boyette, C.. Flint, Michigan: State and and City Agree to Replace 18,000 Home Water Lines by 2020 - CNN. https://edition.cnn.com/2017/03/27/us/flint-settlement/index.html.

51. Smith, M. (2019). Flint Water Prosecutors Drop Criminal Charges, with Plans to Keep Investigating (The New York Times).

52. Carmody, S. (2020). Flint residents may now sue government officials over contaminated water. https://www.npr.org/2020/01/29/800938090/flint-residents-may-now-sue-government-officials-over-contaminated-water.

53. Mason, J., and Messinger, J. (2020). Global gathering to explore how technology is improving forest governance. https://blog.globalforestwatch.org/uncategorized/global-gathering.

54. EU FLEGT Facility. (2020). EU FLEGT Facility: Highlights and Insights from 2017 (FLEGT). http://www.euflegt.efi.int/publications/eu-flegt-facility-highlights-and-insights-from-2017.

55. Ambrose, M.L. (2013). It's about time: privacy, information lifecycles, and the right to Be forgotten. Stan. Tech. L. Rev. *16*, 369–422.

56. Blaser, L. (2014). Old weather: approaching collections from a different angle. In Crowdsourcing Our Cultural Heritage, pp. 45–56.

57. Johnson, J.A. The Value—and Limits—of Distributive Justice in Information Privacy. Paper presented at Western Political Science Association 2016 Annual Meeting, San Diego, CA, 23-26 March.

58. Shirk, J.L., Ballard, H.L., Wilderman, C.C., Phillips, T., Wiggins, A., Jordan, R., McCallie, E., Minarchek, M., Lewenstein, B.V., Krasny, M.E., and Bonney, R. (2012). Public participation in scientific research: a framework for deliberate design. Ecol. Soc. *17*, art29.

59. Evans, B. (2017). J. Power to the people: data citizens in the age of precision medicine. Vanderbilt. J. Entertainment. Technol. *19*, 243.

60. Fasenfest, D. (2019). A neoliberal response to an urban crisis: emergency management in Flint, MI. Crit. Sociol. *45*, 33–47.

61. McKenna, B. (2018). The agony of Flint: poisoned water, racism and the specter of neoliberal fascism. Anthropol. Now *10*, 45–58.

62. Grimmer, C. (2017). Racial microbiopolitics: Flint lead poisoning, detroit water shut offs, and the "matter" of enfleshment. Comparatist *41*, 19–40.

63. Takahashi, B., Adams, E.A., and Nissen, J. (2020). The Flint water crisis: local reporting, community attachment, and environmental justice. Local Environ. *25*, 365–380.

64. ESRI (2016). Mapping Indigenous Territories in Africa (Esri). https://www.esri.com/about/newsroom/arcnews/mapping-indigenous-territories-in-africa/.

65. EU FLEGT Facility. (2020). The Republic of the Congo-EU Voluntary Partnership Agreement (FLEGT). http://www.euflegt.efi.int/background-congo.

66. White, S.C. (1996). Depoliticising development: the uses and abuses of participation. Dev. Pract. *6*, 6–15.

67. Fraser, N. (2009). Scales of Justice: Reimagining Political Space in a Globalizing World (Columbia University Press).

68. Fraser, N.. Reframing Justice in a Globalizing World. https://newleftreview.org/issues/ii36/articles/nancy-fraser-reframing-justice-in-a-globalizing-world.