Published in final edited form as:

Inf Commun Soc. 2019 ; 22(7): 1012–1028. doi:10.1080/1369118X.2019.1596293.

When Data Justice and Environmental Justice Meet: Formulating a Response to Extractive Logic through Environmental Data Justice

Lourdes A. Vera^{a,*}, Dawn Walker^b, Michelle Murphy^c, Becky Mansfield^d, Ladan Mohamed Siad^e, Jessica Ogden^f, EDGI^g

^aSociology and Anthropology, Northeastern University, Boston, MA, USA;

^bFaculty of Information, University of Toronto, Toronto, ON, Canada;

^cHistory and Women and Gender Studies, University of Toronto, Toronto, ON, Canada;

^dGeography, Ohio State University, Columbus, OH, USA;

^eFaculty of Design, Ontario College of Art and Design, Toronto, ON, Canada;

^fSociology, University of Southampton, Southampton, UK;

^gEnvironmental Data and Governance Initiative

Abstract

Environmental data justice (EDJ) emerges from conversations between data justice and environmental justice while identifying the limits and tensions of these lenses. Through a reflexive process of querying our entanglement in non-innocent relations, this paper develops and engages EDJ by examining how it informs the work of the Environmental Data & Governance Initiative (EDGI), a distributed, consensus-based organization that formed in response to the 2016 US presidential election. Through grassroots archiving of data sets, monitoring federal environmental and energy agency websites, and writing rapid-response reports about how federal agencies are being undermined, EDGI mobilises EDJ to challenge the ‘extractive logic’ of current federal environmental policy and data infrastructures. ‘Extractive logic’ disconnects data from

*Please direct all correspondences to Lourdes Vera, l.vera@northeastern.edu.

Notes on contributors:

Lourdes Vera is a PhD Student in Sociology at Northeastern University. She studies environmental data practices for health and justice.

Dawn Walker is a PhD Student at the Faculty of Information. Her research focuses on participation in civic technology and design practices.

Michelle Murphy is Director of the Technoscience Research Unit and Professor of History and Women and Gender Studies at the University of Toronto. Her research focuses on decolonial approaches to environmental and data justice with a focus on the Great Lakes. She is urban Métis from Winnipeg.

Becky Mansfield is Professor of Geography at the Ohio State University. Her research focuses on chemical geographies, environmental health knowledges, and the politics of deregulation.

Ladan Mohamed Siad is a community-supported researcher, designer, and artist. Currently pursuing MDes at Ontario College of Art and Design (OCAD) in the Digital Futures program in Faculty of Design. Ladan’s work focuses on intersections of design, technology, and art, creating multidisciplinary and multisensory interventions surrounding embodied experiences, transnational blackness, data justice, black queer trans feminist frameworks, and human-data interaction.

Jessica Ogden is a PhD Candidate based in Sociology and the Web Science Centre for Doctoral Training at the University of Southampton. Jessica’s research focuses on the politics of data, web archiving and digital data scholarship.

⁴This form of rapid academic analysis is in line with previous EDGI commentary on the ‘Honest and Open New EPA Science Treatment Act’ or HONEST Act (H.R. 1430) (Underhill et al., 2017).

provenance, privileges the matrix of domination, and whitewashes data to generate uncertainty. We use the dynamic EDJ framework to reflect on EDGI's public comment advising against the US Environmental Protection Agency's proposed rule for Transparent Science. Through EDJ, EDGI aspires to create new environmental data infrastructures and practices that are participatory and embody equitable, transparent data care.

Keywords

environmental data justice; extractive logic; whitewashing; data justice; environmental justice

Introduction

The 2016 United States presidential election intensified questions about the integrity, accessibility, and continuity of federal environmental policy and data infrastructures. In response, the Environmental Data & Governance Initiative (EDGI) formed as a distributed, consensus-based organization that seeks to create new communities of practice around documenting, contextualizing, and analysing changes to environmental governance and data practices. EDGI activities include grassroots data archiving; monitoring changes to federal environmental agency websites; interviewing federal Environmental Protection Agency (EPA) and Occupational Health and Safety employees; and writing rapid-response reports about threats to environmental justice, such as the undermining of EPA and important web-based federal environmental resources (Sellers et al., 2017; Paris et al., 2017; Dillon et al., 2017; Underhill et al., 2017; Mansfield et al., 2018).¹ Rooted in critical race, feminist science, environmental justice, and data justice scholarship, EDGI's values of justice, environmental and human health, intersectionality, accessibility, responsivity and proactivity, participatory knowledge making and anti-fascism/racism/oppression (<https://enviroadatagov.org>) serve as the lens of this paper.

EDGI considers environmental data from environmental justice (EJ) and data justice (DJ) perspectives. Early in the formation of EDGI, members realized that while both are grounded in social justice, activism, and action-based engagement, there are also important tensions between EJ and DJ. In response, historian Michelle Murphy offered the term 'Environmental Data Justice' (EDJ) to draw more explicit attention to the challenges of confronting problematic, often industry-produced, data while also striving towards the positive project of building better data practices. Dillon et al. then provisionally defined EDJ as embracing 'public accessibility and continuity of environmental data and research, supported by networked open-source data infrastructure that can be modified, adapted, and supported by local communities' (2017). This preliminary definition of EDJ attended to access, infrastructures, and use, drawing inspiration from ongoing activities of many activist groups working with environmental data, who undertake counter-mapping, community air monitoring, archiving, and tracking of environmental data. From this starting point, we continue to develop the framework of EDJ as an evolving project of reflexively critiquing and repairing the relations of power that always accrue to data as it is made and used; EDJ

¹Currie & Paris, (2018); Lamdan,(2018); Dillon et al., (2017); Walker et al., (2018) discuss the breadth of EDGI's activities.

reaches towards the continuing yet always distant horizon of justice and is an important arena where visions of data justice are being worked through.

To advance an EDJ framework in practice, EDGI members formed an EDJ working group and wrote about EDJ in reports and academic articles (EDGI, 2017). EDGI's report *'Pursuing a Toxic Agenda'* mobilises EDJ to contextualize the politics of evidence around the use and (de)regulation of chlorpyrifos, a harmful toxic pesticide (Paris et al., 2017). Walker et al. (2018) develop EDJ by reflecting on how EDGI's archiving group put EDJ into practice while grappling with the tensions between EJ and DJ. Dillon et al. (2019) draw on feminist science studies to describe how EDJ situates data and moves beyond damage-centred research that documents the harmful effects of environmental violence and data exploitation, instead centring on desire-based research oriented towards building better alternative ways forward (Tuck, 2009; Tuck & Yang, 2016). This paper further develops EDJ by exploring some of its tensions.

Our aim more specifically is to develop an EDJ framework, rooted in EDGI's values, by interrupting what we call an 'extractive logic:' the logic of pulling relations out from bodies and lands into data directed towards other ends. This logic is tied to what Macarena Gomez-Barris (2017) calls the 'extractive view,' which she conceptualizes as the colonial gaze and reorganization of territories for capital accumulation. The extractive logics that animate the creation and use of data are produced by and sustain the matrix of domination, which Black feminist scholar Patricia Hill Collins describes as the interrelated forces of white supremacy, heteropatriarchy, ableism, capitalism, settler colonialism, and other forms of oppression (Collins, 1990; Costanza-Chock, 2018). A key point is that the extractive logic of data haunts any use of data, even when used towards the ends of justice. As an unfinished project, EDJ must constantly reference extractive logic because no data practice can fully escape the pull of non-innocent relations - the embeddedness in racial capitalism and colonialism - that shape data. Acknowledging and learning to refuse this extractive logic through participatory knowledge making and contextualizing data, among other practices, is central to our critical and reflexive EDJ framework.

This paper locates EDJ within currently evolving EJ and DJ scholarship and activism. We use iterative processes of documentation, reflection on practice, and imagining continuing the process of refining an EDJ framework that offers a distinct lens for viewing concerns central to DJ: what informs and who benefits from data-driven decision-making, what counts as data, and what constitutes data ownership. We begin by documenting the practices of activists and scholars and elaborating on extractive logic, which we see as a shared challenge across data and environmental justice. We then reflect on practices that encompass EDGI's internal and external relationships, providing a detailed analysis of EDGI's experience writing a public comment on a proposed EPA regulation for 'transparency in science.' We conclude by imagining an EDJ approach towards environmental data governance that incorporates community perspectives to inform the ontology of environmental data and provide a method of practice and reflection to further enrich understanding.

The conversation between environmental justice and data justice

Both the EJ movement and emerging DJ field strive towards more equitable social change. This section addresses this shared grounding in social justice, activism, and engaged scholarship while surfacing the tensions that arise between EJ and DJ.

Environmental justice scholarship and activism

The EJ movement emerged as a response to the disproportionate exposure to pollutants and environmental hazards borne by marginalized and racialized communities. Reconceiving the environment as ‘where we live, work, and play’ (and eat, learn, and pray), EJ emphasizes the far-reaching effects of structural inequalities on people’s well-being. It is a movement that ‘embraces the principle that all people and communities are entitled to equal protection of environmental and public health laws and regulations’ (Bullard 1996, p. 493). Landmark moments include the 1982 mobilisation by a majority-Black community against a PCB (polychlorinated biphenyl) landfill in Warren County, North Carolina (US EPA, 2015) and the 1987 report ‘Toxic Wastes and Race in the United States,’ which documents the correlation between race and location of toxic waste facilities (United Church of Christ Commission for Racial Justice, 1987). In 1991, the First National People of Color Environmental Leadership Summit adopted 17 principles of EJ that recognized ‘over 500 years of colonization and oppression’ leading to environmental injustices including violence, genocide, and bodily harms (<http://www.ejnet.org/ej/principles.html>). Efforts such as these brought change to US federal laws and policies, including a 1994 Executive Order by President Bill Clinton and subsequent establishment of EPA’s Office of Environmental Justice (US EPA, 2015).

Despite these gains, US environmental policies and federal data infrastructures remain permission-to-pollute systems that perpetuate environmental racism and injustice. Environmental racism implicates the state as complicit in harm, often through insufficient and damaging regulations and laws, lack of funding, relationships with industries responsible for pollution and hazards, or being the party responsible for pollution. In addition, recent changes under the Trump Administration signal a strong de-prioritization of EJ within federal agencies, for example by moving the Office of Compliance and Enforcement Assurance into the Office of Policy (Paris et al. 2017; Bravender, 2017).

Given all this, recent calls have emphasized critical and intersectional approaches to environmental justice frameworks that view environmental racism as part of the matrix of domination and a form of state violence (Ryder, 2017; Pellow, 2016; Voyles, 2015; Liboiron et al.; 2018, Murphy, 2017). Doing so has implications for how EJ addresses questions of knowledge and data. Anthropologist Kim Fortun (2012) has documented how the shifting interactions between environmental groups and information technologies reflect wider changes in views on the production of knowledge, objectivity, and risk. Indeed, EJ often calls for increased data in order to document inequality and environmental racism (Mah, 2017). Further, data to document and try to redress EJ issues is collected through practices such as community monitoring, community mapping, and biomonitoring (of blood, hair, etc.).

However, even when done with good intentions, these are still extractive practices, pulling data about harm and suffering from communities in order to document violence to land and bodies. Too often, EJ involves ‘helicopter science,’ in which scientists enter a community, take samples, and leave; the information is then used for scientific papers and environmental assessments about communities. A different approach — one more aligned with EDJ — is to replace extractive data practices with community-based participatory research and community science projects (Shepard, 2002), in which participants are empowered as fellow researchers and take part in designing the study, collecting data, and analysing it with an emphasis on co-learning (Minkler & Wallerstein, 2008). For example, the Louisiana Bucket Brigade developed inexpensive buckets to enable community members to sample air around polluting facilities (<http://www.labucketbrigade.org/>), a practice that inspired groups across the country to form their own bucket brigades. The increasing popularity of community science has promoted the emergence of online communities such as Public Lab, where researchers, technologists, and community members collaborate to design and use D.I.Y. tools for environmental monitoring (www.publiclab.org).

Data justice’s distinct voices

DJ shares EJ’s concerns around extractive data practices. For instance, companies and governments can easily surveil and predict user behaviour, bringing up questions of data ownership, accessibility, and transparency. DJ offers critical accounts of data-driven surveillance (Dencik, Hintz, & Cable, 2016) and data violence (Hoffmann, 2018), as well as efforts to catalogue data harms (Redden & Brand, 2018). Communities and activists have responded with principles of ‘data justice’ and ‘digital justice’ in an effort to mitigate and engage with these potentially harmful effects of inequitable access to and use of digital technology (<https://www.alliedmedia.org/>).

DJ scholarship has highlighted forms of digital datafication — the increasingly widespread collection of human behaviour data by government and corporations through a process of “life mining” (Schoenberger & Cukier, 2013; Van Dijck 2014). Tech and advertising companies extract insights on personal user data for capital accumulation (Vaidhyanathan 2011), backed by an ideology Van Dijck (2014) calls “dataism” that fosters trust between users and technological platforms. Van Dijck (2014) urges us to examine the roles of corporations, academics, and government in these extractive data practices from an ecosystemic perspective that recognizes the interconnectivity of both ‘operational logic’ and infrastructure. Similarly, Linnet Taylor (2017) uses an ecosystemic ‘capabilities approach’ to justice, bridging three pillars into which DJ scholarship tends to separate: (in)visibility, (dis)engagement, and (anti)discrimination; her approach recognizes the impacts, responses, functionings, and capabilities of data as they are shaped by institutional and political contexts. Feminist approaches to data visualization and frameworks for design justice seek to move beyond damage-centred research to envision new design possibilities. Catherine D’Ignazio and Lauren Klein (2016) bring the tenets of rethinking binaries, embracing pluralism, examining power, considering context, legitimizing embodiment and affect, and making labour visible. And Sasha Costanza-Chock offers an intersectional feminist lens that challenges the matrix of domination to envision a design justice framework that ‘centre[s] the voices of those who are directly impacted’, values expertise based on experiences, and

prioritizes accountability, accessibility, and collaborative processes (2018, p. 2). This scholarship urges us to examine how datafication can be used for emancipation instead of prioritizing capital accumulation.

Activist data projects take an intersectional and liberatory, practice-based approach towards DJ. These groups are oriented towards reconfiguring data infrastructures to incorporate unheard perspectives and move beyond documenting harm. Groups such as Data 4 Black Lives, a ‘group of activists, organizers, and mathematicians committed to the mission of using data science to create concrete and measurable change in the lives of Black people,’ examines and reimagines current data infrastructures for anti-racist and anti-oppressive action (<http://d4bl.org>; Drake 2016). They argue that data systems can empower communities of colour but have historically been used as a tool for discrimination and oppression. Resonating with this critical perspective, the Detroit Digital Justice Coalition formed in 2009 to address the digital divide within Detroit and generate knowledge with community members around use, rights, and ownership of technology. Their principles of ‘digital justice,’ inspired by EJ principles, include access, participation, common ownership, and healthy communities (<https://www.alliedmedia.org/ddjc/principles>). These projects explicitly start from a place of recognition and witnessing of digital racism, as with EJ activists, while further linking it to the aspirational project of articulating principles of a non-racist system that refuses ongoing forms of extraction.

Convergences and tensions across EJ and DJ

Many EJ practices, from community monitoring to environmental assessments, centre on the ethics of data and exemplify EDJ. Both EJ and DJ already recognize the ways in which environmental science and data can perpetuate injustice within the matrix of domination. However, the way they negotiate data collection and understandings of surveillance can sometimes be at odds. DJ scholars flag the myriad and nuanced ways that data collection by the state and private companies function within broader systems of surveillance capitalism (Dencik, Hintz, & Cable, 2016; Cheney-Lippold, 2011) and uphold the matrix of domination (Costanza-Chock, 2018). Grassroots DJ activists not only raise concern over excessive and racist surveillance (Browne, 2015) but also argue for the right to be forgotten. In contrast, most EJ projects call for gathering more data, in the name of a ‘right to know.’ EJ groups often find themselves treating data as a kind of value-neutral object that the state deems acceptable out of necessity *while* questioning the politics of the state and what counts as evidence (Ottinger, 2013). For example, community air monitoring often attempts to gather data about pollution tactically using standards the state will accept, while at the same time critiquing the state for allowing and not monitoring that pollution in first place. Another tension is that increased reliance on digital systems produces an almost compulsory way of life reliant on electronics and devices that in their manufacturing and disposal compound the EJ issues communities face.

Extractive data logic

The act of pulling relations from bodies and lands into data, what we are calling ‘extractive logic,’ is embedded in all types of data. This extraction ignores the situatedness of data,

regarding it as a resource to be pulled out for free, without relations or responsibilities. In its uncritical mode, data is treated as a raw material to be harvested. The dominant framework of corporate, government, and scientific data practices is therefore one of extraction.

Extractive logics are so embedded in data practices with which social justice projects are necessarily, and non-innocently, entangled. For example, within EJ, desire for more data not only demands that harms be proven through a technically narrowed mode of evidence legible to dominant institutions, but also often requires communities to generate evidence of environmental violence from their own suffering. The collection of ‘evidence’ and ‘documentation,’ even under a rubric of transparency and with consent, risks replicating hierarchical structures and the ‘scientism’ of industry and the state without truly addressing or ameliorating the problems communities face.

In these ways, data is both *charismatic and violent*. Data can be used to speak convincingly for complex issues such as climate change or it can be used to more efficiently exclude and marginalize through technological redlining (Thatcher, 2013; Dalton & Thatcher, 2014), predictive policing (Ferguson, 2017), financial scoring (Citron & Pasquale, 2014) and surveillance (Lyon, 2014). Offered up as raw information, data -whether in the form of simple trend graphs or complex machine-learning derived correlations- falsely promises to reveal relations in the world. Gitelman et al. (2013) argue that data is never raw, but always embedded in the politics of its production and interpretation. Helen Kennedy has powerfully illuminated the ‘desire for numbers’ and the affective dimensions of data, data activism, and everyday experiences with it (2016).

Accounting for the range of ways data inhabits a double mode of charisma and violence requires engaging more centrally with the refusals of data violence articulated by indigenous scholars, who have highlighted the way that collecting data on harms continues to place burdens on already targeted, surveilled, harassed, and studied communities. Unangax scholar Eve Tuck (2009) urges a moratorium on these practices of damage-centred research that are built around collecting data about the victimization of marginalized people. M ori scholar Linda Tuhiwai Smith (1999) calls instead for decolonized methodology in which research is guided by Indigenous values.

In response, we propose that EDJ requires tools focused not only on consent to data creation but also refusing the logic of extraction that so profoundly structures data and environmental violences. Our theorization of extractive logic extends from acts of mining and non-consensual taking of land to include myriad practices that commonly make up conventional environmental data produced by the state, companies, and even researchers. These practices – discussed further below – include:

1. Erasure of relations between data and the infrastructures, labour, and lives that produce it.
2. Reproduction of the matrix of domination.
3. Generation of uncertainty through whitewashing.

Disconnecting data from the underground

An extractive logic considers data and other materials as existing in an underground of resources, including but not limited to the earth's underground from where resources such as petroleum, minerals, and coal are extracted (Kinchy et al., 2018). We expand on this concept to consider the 'data underground,' which can take many forms, e.g. a machine learning algorithm that creates data about itself, components of a satellite that will take pictures of the earth, or a community's life activities that are being tracked. These data-producing spaces are proxies for relations and sociotechnical structures. In other words, rather than considering data as raw and extracted from an underground that has existed *a priori*, we see the underground as a politically, economically, culturally, and technoscientifically produced space from where data are extracted. We follow Abby Kinchy et al. (2018) and see the data underground as a proxy for structurally produced sets of often invisibilized labour, practices, and relations that make taking possible.

Practices that put forward 'raw' data erase these relations by not taking into account the full lifecycle and sociotechnical context of the data from its production to dissemination. For example, industry-standard methodologies for data mining, such as the CRISP-DM framework (Wirth & Hip, 2000), position data scientists to decide which variables to exclude, impute missing data, and choose models. Context attributed to the data is often incomplete or only includes the analyst's priorities such as cost-benefit analyses, which in turn are often driven by the material interests of racial capitalism (i.e. ways that capitalism operates through racial difference within processes such as colonization, segregation, slavery, and imperialism (Robinson, 2000; Pulido, 2016)). Even environmental air monitoring pulls information from the air into the possession of companies or the state, mirroring the capitalist extractivism that positions land, air, and water as free resources to be both be mined and wasted. To this, we might add the habitual severing of data mining from the material conditions of its making - from the earthly extractions and labour exploitations that render computational devices possible, to the energy demands of computation itself.

Privileging the matrix of domination

Whether extraction is aimed at the earth or human activity, the extractive view involves a politics of visibility — who gets to see the 'underground' of the earth or, in the case of data, the infrastructures, life supports, and relations that digital practices excavate. Seeing the underground is not just a matter of looking but requires specialized tools and knowledges in order to turn matter and relations into value-making products (Kinchy et al., 2018). Some of these sensing technologies travel between earthly and digital extraction practices. For instance, Light Detection and Ranging (LiDAR) is a commonly used method of remote sensing to determine where to extract energy resources. LiDAR has also aided the construction of a 'virtual wall' along the US-Mexico border (Metz, 2018), tracking the movement of individuals across politically contested landscapes, and continues to be used in the earth sciences to construct 3D representations of rock strata in outcrops.

Gómez-Barris (2017) suggests we consider the 'submerged perspective' when we think about the underground. Often, community and local perspectives are pushed aside in the interest of companies that perpetuate systems of oppression. For instance, the Eviction Lab

has been accused of extractive data practices such as purchasing of eviction data from companies who profit from assisting landlords and failing to involve community groups who have put long-term effort into collecting this data (Aiello et al., 2018). Extractive logics, we contend, tend to rely on and reproduce the matrix of domination, requiring social justice projects — including those oriented towards EDJ — to examine their attempts to mobilise justice through digital technologies.

Whitewashing and spreading uncertainty

Whitewashing is a core strategy for erasing the relations that make up extraction. The act of whitening, covering up, and cursorily excusing through biased or limited engagement has been central to the production of uncertainty around environmental data. In popular culture, whitewashing is seen in the casting of white actors as historically racialized characters in film (Nishime, 2017) and the retouching or modification through makeup of models' skin tone to appear whiter (Pinnock, 2016). We use the term purposefully to highlight interlinked histories of denial of responsibility, curtailed investigation or reliance on biased data, and the production of whiteness.

In the broader context of extractive logics and environmental data, we see whitewashing operate in government data sets, including those commonly relied on by EJ organizers. In particular, the EPA Enforcement and Compliance History Online² and Toxic Release Inventory³ use data that is self-reported by industries, resulting in imprecise and structurally under-reported emissions data (Paris et al., 2017). Moreover, industry produces much toxicological data on chemicals. Regulations are often derived from 'highly technical information brought from experts hired by companies that inform policy,' as well as by government scientists, often pushing aside different knowledge, skills, and resources that activists and local residents can contribute (Ottinger & Zurer, 2011; Baron et al., 2009). This is compounded by the fact that communities typically have no involvement in the design, ownership and possession of environmental data, which instead is held by the perpetrating corporation. Historians of science have documented the gerrymandering of experimental design and purposeful proliferation of divergent scientific findings by industry as part of a concerted project of spreading doubt about environmental harms (Oreskes, 2011; Markowitz, 2013). These regimes of industry and government self-policing around environmental data function to cover up or limit understanding of the relations between the infrastructure creating environmental violence and the data presented in the database.

The structural whitewashing of environmental data makes harm imperceptible, leading many researchers and grassroots activists to strategically extract health data from communities, turning the people carrying the burden of environmental violence into sources of environmental data. In turn, environmental justice and community science groups find themselves uncomfortably relying on this legacy of extractive logics via the practices of community surveys or because they strategically use government and scientifically accepted

²-ECHO contains compliance and enforcement-related information from EPA, spatial, local, and tribal records as well as additional environmental data sets: <https://echo.epa.gov/resources/echo-data/about-the-data>.

³-TRI frames discussion of known issues with self-reported data under the heading 'Data Quality': <https://www.epa.gov/toxics-release-inventory-tri-program/tri-data-quality>.

protocols in their efforts to create legitimated data about harms. Even good faith researchers can reproduce the logics of extractive research, for example enacting a thin version of informed consent and practicing within institutions shaped by a long history of colonial and racial research. This pull to extractive logics is strong, especially because what counts as legitimate forms of environmental data has largely been set by extractive industries such as mining, fossil fuel extraction, and petrochemical production, as well as by states reliant on resource extraction.

Thus, a half century into large-scale environmental data production, we have ended up with a world in which both good faith researchers and bad faith whitewashers extract data from bodies, communities and lands in the forms of health, ecological, and emission data. Any promise of liberation through a critical justice approach to data must contend with this dominant extractive logic that is central to racial capitalism.

EDGI as a place struggling with EDJ in practice

A core methodological commitment within EDGI is the critical analysis of not only state-produced environmental data, but also our own practices. In attempting to design and build positive practices to enact EDJ, EDGI members are encouraged to repeatedly contextualize and analyse their own imbrications in challenging data injustices, including extractive logics. This section of the paper activates this methodological commitment through a self-study of the writing of a public commentary on a proposed EPA rule. We consider our own struggles with extractive logics and the matrix of domination as we strive towards a better vision and practice of EDJ.

In August 2018, EDGI released a public comment critical of EPA's Proposed Rule 'Strengthening Transparency in Regulatory Science' ('Transparent Science' rule) (Mansfield et al., 2018). In appealing language about using the 'best available science' to 'enhanc[e] the public's ability to understand and meaningfully participate in the regulatory process' (Strengthening Transparency in Regulatory Science, 2018), EPA's proposal requires that data and models underlying scientific studies that are pivotal to regulatory action be made available to the public. EPA's short-lived Administrator Scott Pruitt, a Trump appointee, said this rule was designed to ensure that '[t]he science that we use is going to be transparent, it's going to be reproducible, it's going to be analysed by those in the marketplace, and those that watch what we do can make informed decisions about whether we've drawn the proper conclusions' (Kormann, 2018; Friedman, 2018). While it was dubbed a rule against 'Secret Science,' a range of EJ, public health, Congressional, scientific, and even industry representatives offered criticisms and called for the rule to be shelved (Waldman, 2018; Green, 2018; Tollefson, 2018). Even though transparency sounds like a data justice value, critics agree that this rule used the rhetoric of transparency in order to exclude forms of health and community research from regulatory deliberations.

EDGI members drafting this comment came from different disciplines, strengthening our ability to provide both a detailed critique on how the proposed regulation could adversely impact dose response science and the broader social and political implications of limiting data inclusion and availability. In the comment, EDGI calls for an alternative notion of

transparency in regulatory science ‘that incorporates privacy, accessibility, scientific rigor, contextualization, participation, and timeliness’ (Mansfield et al., 2018). Broadly, this form of transparency is not about data in the abstract but instead incorporates how participatory knowledge making includes not only the standard scientific and industry perspectives that have historically shaped environmental regulations, but also the diverse, situated perspectives of multiple marginalized communities. Public and environmental health research has shown that in many cases, the incorporation of participant-centred, local knowledge and values has improved scientific design and rigor (Anderson et al., 2015).

Engaging with an EDJ framework, EDGI’s comment combines EJ’s emphasis on the need for good evidence with DJ’s emphasis on data harms to highlight the rule’s extractive logic. The comment demonstrates that the proposed rule weaponises the language of transparency, not to improve public understanding of the harms of toxic chemicals, but instead to dismantle scientific data as a means to facilitate the weakening of environmental and public health protections and enhance the deregulatory interests of industry. Moreover, equating availability of ‘raw’ data with transparency potentially exposes sensitive data, erasing context-specific concerns about privacy, thus disconnecting data from its provenance. In demanding that all data be available, the rule also requires data dumping of technical analyses, which does not promote public accessibility or meaningful public engagement but instead promotes scientific manipulation – a longstanding tactic of polluting industries trying to avoid regulation. In this way, the EPA’s proposed rule engages in whitewashing as it encourages a tactic that produces more uncertainty about environmental harms.

Scaling down from this federal level critique to EDGI’s organizational practices, an EDJ framework helps us think about the challenges that we faced translating diverse points into concrete, policy-oriented statements that would be heard by the intended audience of regulators. We made a conscious decision not to use words such as “justice” and “precaution.” Not only did we fear that these words would be red flags for some in our audience, but also fully articulating a new, non-extractive vision is difficult within the short timeframe for responding to proposed rules. It was much easier to situate our argument within the extractive logic of cost-benefit analysis that we did not need to justify, showing that the proposed rule does not adequately address the cost of the regulation but makes scientific studies more costly in the long-run.

Similarly, even as we called for contextualized, participatory research in the comment, we did not elaborate on the regular incorporation of indigenous and community knowledges through community and tribal board partnerships. It was much easier to strategically frame the comment around the idea of ‘scientific rigor’ through an EDJ lens but without using EDJ rhetoric. For example, the comment did not explicitly include the concept of ‘strong objectivity’ that feminist philosopher Harding (1992) articulates, but it similarly argues that in order to enhance scientific rigor, data and evidence need to be situated and contextualized to be meaningfully evaluated. On the other hand, neither did the comment use the term ‘biased’ to talk about the intent of the rule, since all knowledge claims are biased in that they stem from particular viewpoints, histories, and blind-spots that need to be acknowledged and situated. A commitment to non-extractive logic requires that we show how, why, and with what effects knowledge is made, shaped and influenced. Along these lines, the commentary

argues that the EPA's proposed rule should be considered relative to the larger political economic backdrop of the use and positioning of data to promote corporate interests over the health of multiple, intersectional marginalized communities. Our EDJ-framed reflection reveals how the commentary was caught by an extractive logic even as it sought to refuse the proposed rule's extractive logic by explicating the 'invisibilized' matrix of domination that undergirds its vision of 'transparent' data.

Enacting EDJ surfaces the tensions produced by extractive logic

Expanding on our reflection on EDGI's recent Transparent Science public comment, we identify three ways that the extractive logic continues to exert influence despite attempts to move beyond it. First, a rhetoric of data 'transparency' and 'openness' does not address fundamental asymmetries of privacy and can counter accessibility and accountability. EDGI's EDJ-framed analysis of the Transparent Science proposed rule shows that ideas about 'transparency' and 'openness' increasingly are taken as uncritically good, yet they can be hollow. As we showed, 'transparency' as deployed in EPA's proposed rule would not improve understanding, but rather encourage tactics like large data dumps of technical data that undermine the ability of scientific data to meaningfully direct environmental protections. There is a similar problem with the concept of 'openness.' Especially under the rubric of Open Data, it has been used to justify construction of data that will more readily cater to corporate interests (Johnson et al., 2017), and has been used to obscure important cases where data is *not* being released (e.g., through facile participation in the Open Government Data initiative) (Mathiesen, 2015). An EDJ frame helps see how governments' movement towards 'open' and 'transparent' data is entangled in an extractive logic, resulting in a disconnection of data from its provenance and from considerations such as usability, ownership, and privacy.

Second, there are many aspects of EDGI activities where attempts to challenge extractive logic have continued to rely upon and replicate it. There is a profound tension in calling out the ways federal environmental data infrastructures whitewash and disconnect data from provenance while simultaneously building sociotechnical systems to ensure that decontextualized federal data remains publicly available (Walker et al., 2018). Yet, we also recognize that as we seek to move beyond damage-centred approaches to justice, many EJ groups still need to document harm in order to achieve timely changes that directly affect health and livelihood. In resisting the dismantling of current environmental protections, EDGI ends up supporting a legacy of government-mandated data collection designed to permit pollution to occur. At the same time, we see that much of the at-risk data we seek to preserve, such as demographic data in conjunction with toxic release inventory reports, offers clear evidence of inequalities. Even the whitewashed data is itself important evidence of the legacy of whitewashing. For this reason, we seek to balance our efforts to enact justice in practice with our broad aspirations for justice.

Third, efforts to create positive politics must carry through critiques of those same politics. As we continue our action-focused projects, which both create participatory civic technologies and new environmental data infrastructures, we look to existing examples of feminist data visualization (D'Ignazio & Klein, 2016) and critical data design practices

(Costanza-Chock, 2018; Fortun et al., 2016) in order to consider how we can deepen our critique of extractive logic, even as it exists in our own practices, through material engagements. Some of these engagements include counter mapping (Kinchy et al., 2018) and community use of satellite and drone technologies to represent local, trans-generational knowledge. These exemplify environmental data practices that challenge the matrix of domination by offering a submerged perspective, which perceives local terrains as the site of knowledge, vitality, and livability (Gómez-Barris, 2017).

Rejecting the overwhelming presence of extractive data practices in racial capitalism and colonialism, EDJ as a framework still sees promise in data as part of justice. A vibrant range of movements has exemplified this refusal, interruption, and alteration of extractive relations towards which EDJ aspires. For example, the Indigenous Data Sovereignty movement advocates for the right of a nation to govern the collection, ownership, and application of its own data (Kukutai & Taylor, 2016). The United States Indigenous Data Sovereignty Network helps ensure that data for and about Indigenous nations and peoples in the US (American Indians, Alaska Natives, and Native Hawaiians) are utilized to advance Indigenous aspirations for collective and individual well-being (Rainie et al., 2017). The broad range and aim of these movements uncover new thinking, emergent practices and techniques that actively combat extractive logics and instead offer robust, community-developed information systems.

Conclusion

Environmental Data Justice as a reflexive framework and practice builds on both EJ and DJ to surface and address tensions between these two, largely complementary approaches to justice. In so doing, EDJ generates ways to critically reflect on how all data practices — including our own — are embroiled in the very extractive logics that we aim to critique and replace. Overall, we want to keep aspiring towards a horizon of justice. We want to uphold a tradition of questioning that informs data-driven decision-making, reimagining what counts as data, and thinking through data ownership. Hence, EDGI's methodological commitment to ongoing documentation, critical reflection, and visioning are crucial to data practice.

Our hope is that EDJ can inform new economic and political infrastructures that shape the ontology of environmental data while challenging extractive logic. For instance, EDJ values such as participatory knowledge making, anti-oppression, intersectionality, and accessibility can be achieved by integrating differently situated knowledges to inform our practices and emphasize feminist practices of care (Dillon et al., 2019). This approach challenges extractive logics by reconnecting data to its provenance through recognition of the perspectives and sociotechnical assemblages that produce or legitimate data. EDJ seeks to un-privilege the matrix of domination by incorporating historically marginalized perspectives into the production, stewardship, and dissemination of environmental data. This participatory knowledge making encourages accessibility and collective ownership of data that troubles the whitewashing and obfuscation of environmental data for corporate interests.

Acknowledging that in addition to experiencing disproportionate environmental impacts, oppressed individuals and communities are further disproportionately burdened with

surveillance and data extraction without consent, we push for environmental data stewardship that takes into account accessibility and transparency while preserving anonymity and confidentiality. We aspire to build chemical inventories contextualized with community experiences and regulations that reflect the acts of environmental violence done to those living adjacent to chemical producing facilities. This would shift the burden of proof from communities experiencing those harms to the industries that are responsible. This contextualization of data can expand to the metadata and re-organization of ‘dumped data’. Open source tools could be flexible and responsive to our critiques as we develop and use them. As we move towards more community-centred governance and regulations, we anticipate being open to critiques of that structure as well. We will continue to elaborate EDJ with acknowledgment of the always unfinished task of building ethical relations and recognizing responsibilities that accrue in the thickness of environmental data.

Acknowledgments

The authors acknowledge the care and labour of EDGI members, in particular the generous conversations within the EDJ working group. This article has benefited from review and feedback from Eric Nost and other EDGI members.

Funding

EDGI has been supported by Doris Duke Charitable Trust and the Packard Foundation as well as individual donors. Research reported in this publication was supported by the National Institute of Environmental Health Sciences of the National Institutes of Health under Award Number T32ES023769. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Aiello D, Bates L, Graziani T, Herring C, Maharawal M, McElroy E, ... Purser G. (2018, 8 22). Eviction Lab misses the mark. Shelterforce. Retrieved from <https://shelterforce.org/2018/08/22/eviction-lab-misses-the-mark/>
- Anderson BE, Naujokas MF, & Suk WA (2015). Interweaving knowledge resources to address complex environmental health challenges. *Environmental Health Perspectives*, 123(11).
- Baron S, Sinclair R, Payne-Sturges D, Phelps J, Zenick H, Collman GW, & O’fallon LR (2009). Partnerships for environmental and occupational justice: Contributions to research, capacity and public health. *American Journal of Public Health*, 99(S3), S517–S525. [PubMed: 19890151]
- boyd, danah, & Crawford K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662–679.
- Bravender R (2017, 9 6). Pruitt tightens political reins on key operations. Retrieved September 27, 2018, from <https://www.eenews.net/stories/1060059803>
- Browne S (2015). *Dark matters: On the surveillance of blackness*. Durham, NC: Duke University Press.
- Bullard RD (1996). Environmental justice: It’s more than waste facility siting. *Social Science Quarterly*, 77(3), 493–499.
- Cheney-Lippold J (2011). A new algorithmic identity: Soft biopolitics and the modulation of control. *Theory, Culture & Society*, 28(6), 164–181.
- Citron DK, & Pasquale F (2014). The scored society: Due process for automated predictions Essay. *Washington Law Review*, 89, 1–34.
- Collins PH (1990). Black feminist thought: Knowledge, consciousness, and the politics of empowerment. Hyman.
- Costanza-Chock Sasha. (2018). Design justice: Towards an intersectional feminist framework for design theory and practice. In *Proceedings of the Design Research Society 2018*. Retrieved from <https://ssrn.com/abstract=3189696>

- Currie ME, & Paris BS (2018). Back-ups for the future: Archival practices for data activism. *Archives and Manuscripts*, 46(2), 124–142.
- Dalton C, & Thatcher J (2014, 5 12). What does a critical data studies look like, and why do we care? Seven points for a critical approach to “big data.” Retrieved September 26, 2018, from <http://societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/>
- Dencik L, Hintz A, & Cable J (2016). Towards data justice? The ambiguity of anti-surveillance resistance in political activism. *Big Data & Society*, 3(2).
- D'Ignazio C, & Klein LF (2016). Feminist data visualization. Presented at the Workshop on Visualization for the Digital Humanities (VIS4DH), Baltimore: IEEE.
- Dillon L, Lave R, Mansfield B, Wylie S, Shapiro N, Chan A, & Murphy M (2019). Situating data in a Trumpian era: The Environmental Data and Governance Initiative. *Annals of the American Association of Geographers*.
- Dillon L, Walker D, Shapiro N, Underhill V, Martenyi M, Wylie S, ... EDGI. (2017). Environmental data justice and the Trump administration: Reflections from the Environmental Data and Governance Initiative. *Environmental Justice*, 10(6), 186–192.
- Drake JM (2016, 6 27). Expanding #ArchivesForBlackLives to traditional archival repositories. Retrieved September 25, 2017, from <https://medium.com/on-archivy/expanding-archivesforblacklives-to-traditional-archival-repositories-b88641e2daf6>
- EDGI. (2017, 9 6). Towards an environmental data justice statement: Initial thoughts. Retrieved September 24, 2018, from <https://enviroidatagov.org/towards-edj-statement/>
- Ferguson AG (2017). *The rise of big data policing: Surveillance, race, and the future of law enforcement*. New York, NY: NYU Press.
- Fortun K (2012). Biopolitics and the informing of environmentalism. *Lively Capital*, 306–328.
- Fortun K, Poirier L, Morgan A, Costelloe-Kuehn B, & Fortun M (2016). Pushback: Critical data designers and pollution politics. *Big Data & Society*, 3(2).
- Friedman L (2018, 4 25). E.P.A. announces a new rule. One likely effect: Less science in policymaking. *The New York Times*. Retrieved from <https://www.nytimes.com/2018/04/24/climate/epa-science-transparency-pruitt.html>
- Gitelman L (Ed.). (2013). *Raw data is an oxymoron*. Cambridge, MA: The MIT Press.
- Gómez-Barris M (2017). *The extractive zone*. Durham, NC: Duke University Press.
- Green M (2018, 6 7). More than 100 bipartisan lawmakers urge Pruitt to scrap “secret science” rule. *The Hill*. Retrieved from <https://thehill.com/policy/energy-environment/391165-bi-partisan-lawmaker-group-calls-pruitt-to-withdraw-secret-science>
- Hardt M, & Negri A (2017). *Assembly*. Oxford University Press.
- Harding S (1992). Rethinking standpoint epistemology: What is “strong objectivity?”. *The Centennial Review*, 36(3), 437–470.
- Hoffmann AL (2018, 4 30). Data violence and how bad engineering choices can damage society. Retrieved September 29, 2018, from <https://medium.com/s/story/data-violence-and-how-bad-engineering-choices-can-damage-society-39e44150e1d4>
- Johnson PA, Sieber R, Scassa T, Stephens M, & Robinson P (2017). The cost(s) of geospatial open data. *Transactions in GIS*, 21(3), 434–445.
- Kennedy H (2016). New data relations and the desire for numbers In *Post, mine, repeat: Social media data mining becomes ordinary* (pp. 221–236). London: Palgrave Macmillan UK.
- Kinchy AJ, Phadke R, & Smith JM (2018). Engaging the underground: An STS field in formation. *Engaging Science, Technology, and Society*, 4, 22.
- Kormann C (2018, 4 26). Scott Pruitt’s crusade against “Secret Science” could be disastrous for public health. *The New Yorker*. Retrieved from <https://www.newyorker.com/science/elements/scott-pruitts-crusade-against-secret-science-could-be-disastrous-for-public-health>
- Kukutai T, & Taylor J (2016). *Indigenous data sovereignty: Toward an agenda* (Vol. 38). Anu Press.
- Lamdan S (2018). Lessons from Datarescue: The Limits of Grassroots Climate Change Data Preservation and the Need for Federal Records Law Reform. *University of Pennsylvania Law Review Online*, 231(166).

- Liboiron M, Tironi M, & Calvillo N (2018). Toxic politics: Acting in a permanently polluted world. *Social Studies of Science*, 48(3), 331–349. [PubMed: 29963976]
- Lyon D (2014). Surveillance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data & Society*, 1(2).
- Mah A (2017). Environmental justice in the age of big data: Challenging toxic blind spots of voice, speed, and expertise. *Environmental Sociology*, 3(2), 122–133.
- Mansfield B, Gehrke G, Hindman A, Fredrickson L, Vera L, Ogden J, & Walker D (2018). EPA'S proposed rule uses the idea of transparency to reduce real transparency and delay protecting environmental and public health (Public Comment) Retrieved from <https://envirodatagov.org/wp-content/uploads/2018/09/Strengthening-Transparency-in-Regulatory-Science.pdf>
- Markowitz G, & Rosner D (2013). Deceit and denial: The deadly politics of industrial pollution (Vol. 6). Oakland, CA: UC Press.
- Mathiesen K (2015). Transparency for democracy: The case of open government data In Moore AD (Ed.), *Privacy, Security and Accountability: Ethics, Law and Policy*. Rochester, NY: Rowman & Littlefield.
- Metz C (2018, 9 18). Parts of a 'virtual' border wall, built with the tech behind driverless cars. The New York Times. Retrieved from <https://www.nytimes.com/2018/09/18/technology/trump-border-wall-lidar.html>
- Minkler M, & Wallerstein N (2008). *Community-based participatory research: From process to outcomes*. San Francisco, CA: Jossey-Bess.
- Murphy M (2017). *The economization of life*. Durham, NC: Duke University Press.
- Nishime L (2017). Whitewashing yellow futures in ex machina, cloud atlas, and advantageous: Gender, labor, and technology in sci-fi film. *Journal of Asian American Studies*, 20(1), 29–49.
- Oreskes N, & Conway EM (2011). *Merchants of doubt: How a handful of scientists obscured the truth on issues from tobacco smoke to global warming*. Bloomsbury Publishing USA.
- Ottinger G (2013). *Refining expertise: How responsible engineers subvert environmental justice challenges*. New York, NY: NYU Press.
- Ottinger G, & Zurer R (2011). Drowning in data. *Issues in Science and Technology*, 27(3), 71–82.
- Paris B, Dillon L, Pierre J, Pasquetto IV, Marquez E, Wylie S, ... EDGI. (2017). Pursuing a toxic agenda: Environmental injustice in the early Trump administration (The First 100 Days and Counting). Environmental Data & Governance Initiative. Retrieved from <https://envirodatagov.org/publication/pursuing-toxic-agenda/>
- Pellow DN (2016). Toward a critical environmental justice studies: Black lives matter as an environmental justice challenge. *Du Bois Review: Social Science Research on Race*, 13(2), 221–236.
- Pinnock O (2016, 9 2). White-washed runways: The effects of racism in the fashion industry. Retrieved September 29, 2018, from <http://www.thefashionlaw.com/home/is-racism-stifling-creativity-in-the-fashion-industry>
- Pulido L (2016). Flint, Environmental Racism, and Racial Capitalism. *Capitalism Nature Socialism*, 27(3), 1–16.
- Rainie SC, Schultz JL, Briggs E, Riggs P, & Palmanteer-Holder NL (2017). Data as a strategic resource: Self-determination, governance, and the data challenge for Indigenous nations in the United States. *The International Indigenous Policy Journal*, 8(2).
- Redden J, & Brand J (2018). Data harm record. Retrieved September 29, 2018, from <https://datajusticelab.org/data-harm-record/>.
- Robinson CJ (2000). *Black Marxism: The making of the Black radical tradition*. University of North Carolina Press.
- Ryder SS (2017). A bridge to challenging environmental inequality: Intersectionality, environmental justice, and disaster vulnerability. *Social Thought and Research*, 34, 85–115.
- Sellers C, Dillon L, Ohayon JL, Shapiro N, Sullivan M, Amoss C, ... EDGI. (2017). The EPA under siege: Trump's assault in history and testimony (The First 100 Days and Counting). Environmental Data & Governance Initiative. Retrieved from <https://envirodatagov.org/publication/the-epa-under-siege/>

- Shepard PM (2002). Advancing environmental justice through community-based participatory research. *Environmental Health Perspectives*, 110(Suppl 2), 139. [PubMed: 11836141]
- Smith LT (1999). Decolonizing methodologies: Research and indigenous peoples. *Strengthening Transparency in Regulatory Science*, Pub. L. No. 83 Fed. Reg. 18768, 40 CFR 30 18768 (2018). Retrieved from <https://www.federalregister.gov/documents/2018/04/30/2018-09078/strengthening-transparency-in-regulatory-science>
- Taylor L (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, 4(2).
- Thatcher J (2013). Avoiding the ghetto through hope and fear: An analysis of immanent technology using ideal types. *GeoJournal*, 78(6), 967–980.
- Tollefson J (2018, 6 1). US EPA science advisers question ‘secret science’ rule on data transparency. *Nature*.
- Tuck E (2009). Suspending damage: A letter to communities. *Harvard Educational Review*, 79(3), 409–428.
- Tuck Eve, & Yang K. Wayne. (2016). What justice wants. *Critical Ethnic Studies*, 2(2), 1.
- Underhill V, Martenyi M, Lamdan S, & Bergman A (2017). Public protections under threat at the EPA: Examining safeguards and programs that would have been blocked by H.R. 1430 (White Paper). Environmental Data & Governance Initiative. Retrieved from <https://envirodatagov.org/publication/public-protections-under-threat/>
- UCC Commission for Racial Justice. (1987). Toxic wastes and race in the United States: A national report on the racial and socio-economic characteristics of communities with hazardous waste sites. New York, NY Retrieved from <http://uccfiles.com/pdf/ToxicWastes&Race.pdf>
- US EPA. (2015, April 15). Environmental Justice. Retrieved September 27, 2018, from <https://www.epa.gov/environmentaljustice>
- Van Dijck J (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197–208.
- Voyles TB (2015). *Wastelanding: Legacies of uranium mining in Navajo country*. University of Minnesota Press.
- Waldman S (2018, 5 16). EPA’s own advisory board questions “Secret Science” plan. *Scientific American*. Retrieved from <https://www.scientificamerican.com/article/epas-own-advisory-board-questions-secret-science-plan/>
- Walker D, Nost E, Dillon L, Lave R, & Lemelin A (2018). Practicing environmental data justice: From DataRescue to Data Together. *Geo: Geography and Environment*.
- Wirth R, & Hipp J (2000). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining* (pp. 29–39).