

December 14, 2023

Dear Dr. Fedinick and the OSTP Environmental Justice Working Group,

We at the Environmental Data and Governance Initiative (EDGI) are grateful for the opportunity to share our experiences and analyses to support the OSTP in developing the Environmental Justice Science, Data, and Research Plan. EDGI is a multidisciplinary and cross-professional network focused on advancing the environmental right to know and the democratic integrity of environmental governance. We are a national leader in tracking and analyzing federal public environmental information, environmental enforcement data, and changes to environmental governance structures. The comments we submit derive from analyses we've conducted over the past seven years. We structure our comments as brief responses to select components of the information requested by the OSTP, and we would welcome further conversation about any and all of the issues raised here.

#### 1. Development and Use of Science, Data and Research to Support Environmental Justice Policy

*(b) What are the biggest opportunities for advancing research and development to support environmental justice-related decision making, both within the Federal research programs and in Federal extramural grant programs?*

We think that two of the biggest opportunities for advancing federal research to support environmental justice-related decision making relate to cumulative exposures and cumulative impacts. These issues are absolutely critical to ending the legacy of environmental injustices in this country, and while the importance of these issues has been acknowledged for many years, robust federally-funded research to compel the federal government to address them has been lacking. Each of these are complex issues, and the federal government's attempts to shoehorn them into a highly constrained risk assessment framework has been unproductive.<sup>1</sup>

---

<sup>1</sup> [https://envirodatagov.org/wp-content/uploads/2023/09/EDGI\\_PublicComment\\_EPA\\_CRA.pdf](https://envirodatagov.org/wp-content/uploads/2023/09/EDGI_PublicComment_EPA_CRA.pdf)

We strongly encourage fundamental research that utilizes recent advancements in exposure science technology as well as social science modeling methods to move these fields forward. We specifically recommend extensive monitoring of hazardous air pollutants and other airborne toxic chemicals to ground truth estimates of facility emissions as well as neighborhood exposures to toxic chemicals.

In order to meaningfully incorporate these exposures and non-chemical stressors into cumulative impacts assessments, we recommend pursuing cutting-edge health justice research involving longitudinal and intergenerational studies to examine temporally compounding impacts of chemical and non-chemical stressors. We further recommend supporting the development of nationally representative studies that build and test observable metrics for latent constructs such as racism, classism and criminalization, such that these constructs can be meaningfully incorporated into cumulative impacts analyses. It is not race, but structural and individual racism, that produce health burdens for people of color in this country and contribute to the cumulative impacts of environmental injustices.

*(d) What data sources should the Environmental Justice Subcommittee consider recommending to the Chair of CEQ for inclusion in the Climate and Economic Justice Screening Tool established pursuant to section 222(a) of Executive Order 14008?*

EDGI strongly recommends the inclusion of environmental enforcement and compliance data, industrial sector data, and historical data in the CEJST tool. The unequal enforcement of existing laws is a fundamental component of environmental racism that must be acknowledged and addressed, and the patterns of violating industries must also be understood in order to address some of the root causes of uneven pollution exposures. Likewise, the importance of cumulative exposures and impacts demand data that have temporal depth - such data may include shapefiles on redlining or composites of pollution or environmental protection law violation rates over time.

In April 2022, EDGI submitted a public comment regarding these issues.<sup>2</sup> Drawing on EPA's Enforcement and Compliance History Online (ECHO) database, we analyzed twenty years of Clean Water Act, Clean Air Act, and Resource Conservation and Recovery Act violations by facilities across the nation. We were able to identify an additional 107 Census tracts that

---

<sup>2</sup> <https://envirodatagov.org/wp-content/uploads/2022/04/CEQ-CEJST-Public-Comment-by-EDGI.pdf>

could be deemed “disadvantaged” - tracts where violations per facility were in the 90th percentile or above, or facility inspections were in the 10th percentile or below, and the CEJST socio-economic criteria were met.

EDGI’s analysis of the draft version of CEJST also identified that fossil fuel industries—especially power plants and pipelines—were present in nearly 80% of the tracts CEJST named as disadvantaged. Identifying and visualizing these patterns is critical for developing regulatory accountability and addressing environmental injustices. CEJST could draw on North American Industrial Classification System (NAICS) codes in metrics such as “Fossil Fuel Industry Concentration” to enable the public to understand the economic forces that create conditions for “disadvantage” across tracts. This would also aid the public and decision-makers in identifying specific solutions to fund, such as community-led energy transition projects.

CEJST only utilizes contemporary measures of environmental harm, such as 2020 wastewater discharge rates in a Census tract. Given the significance of pollution exposures that accumulate over time, CEJST could draw on EPA’s extensive if not perfect archive of reported facility discharges and emissions in order to craft composite metrics like “Annual Average Wastewater Discharge 2000-2020.”

*(f) Please provide examples of data, research, local or Indigenous Knowledge, and/or science—or the lack thereof—that have been misused or misinterpreted in environmental justice-related decisions and actions.*

The lack of data identifying a problem is often misinterpreted as a demonstration that there is not a problem. The Environmental Protection Agency’s entire approach to hazardous air pollutants is subject to this misinterpretation based on a lack of data. There are scant requirements—non-existent in most states—to report actual measurements of hazardous air pollutants emitted from facilities or in ambient air in neighboring communities. Instead, emissions estimates based on facility production estimates are all that is required, leaving gaping holes in agencies’ and the public’s abilities to know what is in the air workers and communities are breathing.<sup>3</sup> Many states do not even own the equipment necessary to monitor toxic chemicals that they regulate. The same goes for the thousands of chemicals that are not required to be reported under the Clean Water Act or

---

<sup>3</sup> <https://www.propublica.org/article/whats-polluting-the-air-not-even-the-epa-can-say>

Safe Drinking Water Act. The demonstration of prevalence and toxicity of chemicals like PFAS in recent years provides a powerful example of how problematic the government's approach to chemical monitoring is. EDGI recommends a vastly expanded chemical monitoring program in order to fill these pervasive chemical exposure data gaps.

We also recommend OSTP develop guidelines cautioning agencies against inappropriate uses of data or the lack thereof in environmental justice analyses. In 2021, we worked with the GreenRoots organization in Chelsea, MA as they sought to contest EPA's re-permitting of several oil storage facilities under the National Pollutant Discharge Elimination System of the Clean Water Act. Unfortunately, the EPA has stated that "the CWA does not appear to provide [the agency] with any general authority to impose permit conditions based on EJ considerations that are not connected to water quality impacts or technology-based limitations."<sup>4</sup> This poses a significant problem from a cumulative impacts perspective because it hinders the agency's ability to regulate based factors beyond the relatively narrow confines of smokestacks, fencelines, and discharge pipes. It also raises questions about why the EPA even conducts EJ analyses if it isn't able to act on them. In the analysis of the Chelsea Creek permits, the EPA documented existing environmental injustices, but was unable to address or incorporate them "because the permit decisions are limited to whether water discharges from the facilities will contribute negatively to water quality standards."<sup>5</sup> Inactionable research wastes agency as well as community time and resources, and stigmatizes communities as "disadvantaged" without providing them any means of redress. Even when the EPA focused its assessment on water discharges and standards, it claimed, without evidence, that the permits "will not have a disproportionately high and adverse human health or environmental effect on minority or low-income populations near the permitted facilities." The EPA's assessment assumed that existing water discharge limits would lead to achieving water quality standards and adequately protecting human health. However, the EPA failed to consider the data showing that these facilities were routinely out of compliance with their permits and exceeded discharge limits. While, if complied with, a certain discharge limit may ensure water quality standards, the EPA should account for the likelihood of non-compliance in setting these limits.

---

<sup>4</sup> <https://envirodatagov.org/edgis-public-comment-on-epas-npdes-renewal-for-chelsea-creek/>. The factsheet with the EPA's quote is no longer available online, but the context and quotation are referenced in our public comment on the issue.

<sup>5</sup> Ibid.

In this example, the EPA ignored available evidence (on non-compliance and on social stressors and vulnerabilities) and made significant misinterpretations of other data (by assuming the facility would comply with permit levels). As such, the EPA's current approach to EJ analyses seems to perpetuate unproductive and potentially harmful research practices that extract valuable time and effort from communities.

## 2. Identifying and Addressing Data Gaps and Inadequacies in Data Collection and Scientific Research Related to Environmental Justice

*(a) What data gaps or data collection challenges have you encountered related to patterns of historical or ongoing discrimination and bias (e.g., related to income, race, color, national origin, Tribal affiliation, or disability)?*

There are several pervasive data gaps that impede analyses of and action on environmental justice, in two broad types: data that are not required to be collected, and data that are required but are not reported anyway. Of the data that are not required to be collected, there is some very basic data, such as measured rather than estimated emissions, and there are also more methodologically and jurisdictionally challenging to acquire, but extremely important, data.

The most basic data for which extraordinary data gaps exist are real-time and aggregate measured emissions and ambient air data. The standard practice is to base decisions on industry-reported estimates of emissions and downstream models of community exposure, rather than empirical measurements. Facilities that estimate they will emit below a certain (very high, 10 tons per year) threshold, are not even required to estimate their annual emissions of hazardous air pollutants. Other common data gaps include the health effects of exposure to multiple pollutants at single time-points as well as health effects of chronic exposure to multiple pollutants, which are essential to understanding cumulative exposures.

There are substantial data gaps surrounding intersectional justice issues. For example, there is a dearth of health and epigenetic data regarding the consequences of living in communities that are hyper-surveilled by the criminal justice system, though health

outcomes for those adjacent to the carceral system are extraordinarily poor.<sup>6</sup> There are still substantial data gaps surrounding the physical and biological consequences of poverty. There are data gaps at the intersection of occupational and environmental health as well, such as compounding physical and biological consequences of exposure to extreme weather. In order to address these data gaps that impede understanding and action, we need more and better data about exposure to systematic discrimination of many kinds.

In addition to data that is not required to be collected, there are remarkably large gaps in data that the federal government already expects to have. For example, in a recent (as of yet unpublished) analysis of carceral facilities' violations of environmental laws, researchers found that only approximately 10% of such facilities were categorized into an industrial sector (NAICS) code, meaning that statistically defensible industry-wide analyses are nearly impossible with existing federal datasets. EDGI has found striking amounts of missing data for all of the major environmental laws.<sup>7</sup> For example, we found that in the ECHO database, more than 19,000 facilities are missing location data, 78% of facilities regulated by the Clean Water Act are missing inspection count data, and 83% of facilities regulated by the Safe Drinking Water Act are missing race data for the surrounding area (percent minority-identified people in a three mile radius), which is essential for environmental justice analyses. There are several cases where data missingness is more pronounced in majority-minority areas. Overall, the extent of missing data precludes analyses at scale and severely inhibits environmental justice-related decision making.

### 3. Encouraging Participatory Science and Meaningful Engagement for Communities

*(a) What role should the Federal government play in collecting, storing, and managing community-derived data, including information collected from communities with environmental justice concerns?*

We believe a paradigm shift with respect to community data is necessary to move the needle on environmental justice.<sup>8</sup> EDGI recommends that the Federal government take an active role in equipping communities to collect and contribute data for environmental

---

<sup>6</sup> <https://doi.org/10.2307/j.ctv2vt05hh>

<sup>7</sup> [https://envirodatagov.org/wp-content/uploads/2022/09/Gaps\\_and\\_Disparities\\_Report.pdf](https://envirodatagov.org/wp-content/uploads/2022/09/Gaps_and_Disparities_Report.pdf)

<sup>8</sup> <https://envirodatagov.org/a-green-new-deal-for-environmental-data/>

governance, managing data at regional and national scales, and developing accessible data portals. A major hindrance for community-derived data is the isolation of that data due to capacity constraints of small organizations, and the Federal government could support communities in providing infrastructure to support the collection, aggregation, and dissemination of community-based data. Moreover, the Federal government must create avenues for community-driven and community-derived data to trigger Federal inquiry and action.

*(b) What suggestions do you have for use of community-derived data in Federal decisions with varying needs for quality assurance, reproducibility, and peer review across different decision contexts?*

There are actions that can be taken commensurate with the quality assurance practices of data received, and there also needs to be scaffolding to escalate an issue from community-derived data to agency-derived data that meets QA/QC requirements. For example, community-derived data suggesting high concentrations of emitted pollutants could and should be followed up with an official site visit. Community-derived studies indicating problematic chemical exposures should also trigger agency studies, such as the case in Tonawanda, NY where concerned residents collected air samples, and the EPA followed up with an investigation that confirmed extremely high toxic emissions and fraudulent records alterations.<sup>9</sup>

*(c) What are the priority decision contexts in which community-derived data should be applied?*

Permit decisions are the priority decision contexts in which community-derived data should be applied. Communities near facilities have lived expertise – they know the patterns of odors, noises, neighbors’ health symptoms. They know the wind patterns and when the fog comes. Currently, permit decisions are almost exclusively determined by industry self-reported estimates of emissions, which are often based on outdated models that lack evidence for accuracy and applicability to different contexts. Community-derived data should be treated with at least the same respect as invalidated industry self-reported data and estimates.

*(f) What practices could ensure that effective, respectful, and meaningful public engagement is built into the research process?*

---

<sup>9</sup> <https://www.cacwny.org/campaigns/tonawanda/>

The key to ensuring effective, respectful, and meaningful public engagement is ensuring the ability of the public to actually influence research questions, design, interpretation, and ultimately decision-making related to the research. Early engagement is essential. The public—community members, academics, non-profits—should be respected partners in setting the research plan, not only brought on after specific budgets are already set and the litany of Federal constraints are in place. Initial engagement could involve listening sessions or solicitation of public partners throughout the process, depending on the issue at hand. Sustained engagement could involve members of the public participating in a workgroup or on a project-specific advisory council. Crucially, public engagement should continue through the use of data and decisions regarding next steps, such as follow-up studies or policy changes.

In an earlier section, we described an example (Chelsea Creek, MA) where public engagement in a permitting decision was not particularly meaningful because of the way the EPA and other agencies are constrained in what factors they can consider in their analyses of environmental justice. We also note a couple of small but significant technical hurdles. First, agencies need to make relevant data actually useful for communities by disseminating it in appropriate formats (such as CSVs rather than tables embedded in PDFs). Second, agencies also need to ensure that they have fully disclosed public comments for review.<sup>10</sup>

One potentially valuable model for engagement stems from our facilitation of a focus group with community organizations where we use-tested EPA’s ECHO Notify alert tool.<sup>11</sup> That tool is meant to help any member of the public track reported violations and enforcement actions at industrial facilities. The EPA contracted with us to facilitate feedback in this way because paperwork reduction rules limit their data collection activities. Through this focus group approach, with an intervening or facilitating organization such as EDGI, community groups would be able to provide key recommendations on data and research tools directly to the EPA.

---

<sup>10</sup>

[https://envirodatagov.org/wp-content/uploads/2023/10/Wylie\\_Northeastern\\_Class\\_Comment\\_EPAREGION1\\_ChelseaOilPermits.pdf](https://envirodatagov.org/wp-content/uploads/2023/10/Wylie_Northeastern_Class_Comment_EPAREGION1_ChelseaOilPermits.pdf)

<sup>11</sup> <https://docs.google.com/document/d/1EDAndZBnkS0hJ5aLliX9KdD8-NndqHFACGcyo8X8zBc/edit>



## 5. Research Coordination and Public Access to Federal Data

*(a) Are there datasets not owned by the Federal government that you have utilized to help support the advancement of environmental justice? If you have used non-Federal data sets to advance environmental justice, which ones have you used and why?*

EDGI has used a variety of non-Federal data sets in our analyses, including (but not limited to) toxic chemical data made available by the non-profit FracFocus,<sup>12</sup> governance data made available by ToxicDocs,<sup>13</sup> community water systems boundaries data made available by the Environmental Policy Innovation Center,<sup>14</sup> and lead service line data made available by the New Jersey Department of Environmental Protection.<sup>15</sup> We have also created our own datasets of changes to federally provisioned public information on environmental topics.<sup>16</sup> We have used these datasets because they are high quality and fill essential gaps in Federal data.

*(c) What kinds of tools and resources would help communities and local decision-makers better access data and information and address environmental justice in decision making?*

A pervasive hurdle to community use of Federal data is the disjuncture between the scope and scale of interest to communities and that which is made available by the Federal government. Often the Federal government reports environmental data from a single facility, but offers no way to aggregate or synthesize data at meaningful geographical units such as Congressional Districts or watersheds. EDGI has demonstrated the power of doing this through our Environmental Enforcement Watch program, and has prototypes the Federal government could adopt.<sup>17</sup> Agencies could produce open source libraries that allow the public to aggregate and/or filter data in order to build analyses of interest to specific communities.

## 6. Data Analysis and Methodological Considerations

---

<sup>12</sup> <https://www.fracfocus.org>

<sup>13</sup> <https://www.toxicdocs.org>

<sup>14</sup> <https://www.policyinnovation.org/technology/water-utility-service-area-boundaries>

<sup>15</sup> [https://geo.btaa.org/catalog/00e7ff046ddb4302abe7b49b2ddee07e\\_13](https://geo.btaa.org/catalog/00e7ff046ddb4302abe7b49b2ddee07e_13)

<sup>16</sup> <https://envirodatagov.org/federal-environmental-web-tracker-about-page/>

<sup>17</sup> <https://environmentalenforcementwatch.org/>

*(b) What methods do you recommend for analyzing cumulative impacts (including risks) from multiple sources, pollutants or chemicals, and exposure pathways, and accounting for non-chemical stressors and current and anticipated climate change?*

Analyzing cumulative impacts is a multidisciplinary, complex challenge that will likely evolve over the coming years as exposure science and modeling become ever more sophisticated. EDGI recommends expanding multi-pollutant exposure studies to assess the end points of exposure to industry-relevant chemical mixtures, as well as population-based health outcome analyses and ambient principal components analyses. We recommend incorporating studies that assess the effects of exposure over varying temporalities. We recommend multiple exposure pathway scenarios incorporating anticipated effects of climate change on air dispersion, water cycling, and permafrost reduction, as well as expanded hazard assessments involving extreme weather events. To include non-chemical stressor effects, we recommend supporting health justice studies examining the biological effects of those stressors and identifying observable items, and then utilizing those observable variables in latent construct analyses. Each of these recommendations points to areas with current cutting edge research that should be supported and incorporated by the Federal government. We recommend developing a Cumulative Impacts Advisory Committee to support the Federal government in designing appropriate methods and analyses.

Sincerely,

The Environmental Data and Governance Initiative (EDGI)