MITRE's Response to the OSTP RFI on Equitable Data Engagement and Accountability

October 3, 2022

For additional information about this response, please contact:
Duane Blackburn
Center for Data-Driven Policy
The MITRE Corporation
7596 Colshire Drive
McLean, VA 22102-7539

policy@mitre.org
(434) 964-5023

<<This page is intentionally blank.>>

About MITRE

MITRE is a not-for-profit company that works in the public interest to tackle difficult problems that challenge the safety, stability, security, and well-being of our nation. We operate multiple federally funded research and development centers (FFRDCs); support public-private partnerships across national security and civilian agency missions; and maintain an independent technology research program in areas such as artificial intelligence, intuitive data science, quantum information science, health informatics, policy and economic expertise, trustworthy autonomy, cyber threat sharing, and cyber resilience. MITRE's 9,000-plus employees work in the public interest to solve problems for a safer world, with scientific integrity being fundamental to our existence. We are prohibited from lobbying, do not develop or sell products, have no owners or shareholders, and do not compete with industry. Our multidisciplinary teams (including engineers, scientists, data analysts, organizational change specialists, policy professionals, and more) are thus free to dig into problems from all angles, with no political or commercial pressures to influence our decision making, technical findings, or policy recommendations.

Good data is essential for making well-informed decisions, managing organizational costs and benefits, innovating, and successfully achieving mission outcomes—which includes ensuring that "government programs and policies yield consistently fair, just, and impartial treatment of all individuals."[1] MITRE not only supports Chief Data Officers throughout the federal government as they develop and implement their data strategies and governance programs, but also has direct experience assisting federal agencies leverage other government and private-sector data to meet critical mission needs. Per the Federal Acquisition Regulation, FFRDCs can have unique access to both sensitive government data and proprietary private-sector data—and both the government and the public sector have regularly trusted MITRE to manage third-party access and leveraging of their data. Multi-party collaboration on ethical data collection, protection, access, and usage is fundamental to our work in this space.

Introduction and Overarching Recommendations

An overarching observation from this RFI and the administration's prior document, *A Vision for Equitable Data,* is the assumption that equitable data has already been collected and efforts can thus be predominantly focused on access and use. In practice, however, these aspects aren't isolated events, but are rather interconnected aspects that both follow the collect-access-use progression and contain feedback loops such that insights gained from access and use will impact future collection activities. As such, MITRE recommends that the community's work on this equitable data topic be more holistically scoped and coordinated to drive the impacts desired.

---

[1] A Vision for Equitable Data: Recommendations from the Equitable Data Working Group. https://www.whitehouse.gov/wp-content/uploads/2022/04/eo13985-vision-for-equitable-data.pdf

Questions Posed in the RFI

1. What are examples of successful collaborations involving equitable data between the Federal government and (a) Tribal, territorial, local, and State governments, or (b) local communities?

This Administration encouraged the use of American Rescue Plan funds to support Tribal, territorial, local, and state (TTLS) data infrastructure and expanding reporting requirements for disaggregated data collection across key programs as a part of funding activities to institutionalize expectations for allocating resources and assessing equity at the state and local levels. "Agencies and program offices will need to invest in the statistical, evaluation, and data science expertise necessary to design and conduct robust equity assessments using their administrative data, consistent with the Foundations for Evidence-Based Policymaking Act of 2018 ('Evidence Act') and the Equity EO."[2] TTLS agencies can collaborate on equitable data through policies and standards that encourage transparency in allocation of resources.

Related lessons learned from the administrative equity assessment and evaluation activities undertaken by state and local levels on equitable data include:[3]

- Establish policies and standards for administering funding and supporting administrative costs to design equitable project data collection, evaluation, and analysis systems and structures that target underserved communities and create opportunities for those who have been systematically excluded in the past.

- Support program evaluation and evidence-based resourcing, such as: data analysis resources to gather, assess, and share equity data; develop shared equitable data infrastructure; encourage community outreach and engagement activities for TTLS governments.

State data centers (SDCs) are the official representatives for the Census State Data Center Program.[4] The value added by SDC members includes enhancing Census data with other information, such as economic statistics or population forecasts, which may be produced in-house or come from sources other than the Bureau. Combining and interpreting Census data with other key measures is a value SDC members provide to the public. Census data helps inform funding and community programming and can be used to inform business strategy, support advocacy efforts, and better understand neighbor characteristics and resources, including equitable data such as race, age, and sex. The U.S. Commission on Civil Rights has found that "data collection and reporting are essential to effective civil rights enforcement, and that a lack of effective civil rights data collection is problematic."[5] SDCs can support community, researchers, advocates, and policymakers with access to disaggregate data and identify disparate impacts of government programs and services for diverse population groups.

---

[2] Ibid.

[3] Ibid.

[4] State Data Center Clearinghouse. 2022. Census Bureau, https://www.census.gov/about/partners/sdc/related-sites.html. Last accessed September 28, 2022.

[5] Are Rights a Reality? 2019. U.S. Commission on Civil Rights, https://www.usccr.gov/files/pubs/2019/11-21-Are-Rights-a-Reality.pdf.

2. Among examples of existing Federal collaborations with (a) Tribal, territorial, local, and State governments or (b) local communities involving equitable data, what lessons or best practices have been learned from such collaborations?

Lesson Learned: Adding lived experience narratives in equitable data collections is critical.

The Treasury's Community Development Financial Institutions (CDFI) fund "plays an important role in generating economic growth and opportunity in some of our nation's most distressed communities."[6] To help support overall program investment and management activities, the CDFI fund has been collecting equitable data from several sources, including the U.S. Census, Federal Reserve Economic Data, Statistics of U.S. Businesses, the Consumer Financial Protection Bureau, Esri, and the U.S. Department of Agriculture Economic Research Service. The CDFI fund is using this equitable data to better understand the distribution of disability concentrations, SNAP benefits, unemployment, household burden, subprime credit populations, housing prices, food deserts, and housing availability for low-income communities.

To support the Treasury's Learning Agenda[7] goal to assess the impact of the CDFI investments to alleviate poverty, an open question emerged around whether existing equitable data provided sufficient causal insights into the factors driving persistent poverty. MITRE's analysis found that the diversity of communities is shaped by a variety of formative experiences and distinct contemporary environments. This analysis hints at complexities beyond the underlying patterns of poverty, which are not fully appreciable by purely statistical methodologies and existing equitable data sources. MITRE found that additional information from the local communities that were the ultimate recipients of these funds clarified the analysis. See Appendix Figure 1 in the appendix for a categorization of additional equitable data elements derived from qualitative interviews.

Lesson Learned: Predictive models incorporating community voices enable proactive interventions to address equity issues before they emerge or are exacerbated.

MITRE further analyzed this issue for CDFI and created a four-step model for deriving insights from an inductive "bottom-up" approach (see Appendix Figure 2 in the appendix). This includes designing participatory and qualitative research methods to include the voices of people with low income and the realities (barriers) of their lives into CDFI research, as well as developing data dashboards to explore poverty patterns and anticipating and proactively supporting downward trending communities before they become PPCs. A similar, but more generic, model could be developed for broader equitable data collections.

Lesson Learned: Usable and accessible visualizations/dashboards enable more, and more diverse, stakeholders to engage with data.

Uploading equitable data into accessible, interactive, web-based dashboards can help bring both the federal and local communities together under a common umbrella. This helps:

---

[6] What Does the CDFI Fund Do? 2022. Department of Treasury, https://www.cdfifund.gov/. Last accessed September 23, 2022.

[7] FY 2022-2026 Department of the Treasury Learning Agenda. 2021. Department of the Treasury, https://home.treasury.gov/system/files/266/Treasury-FY2022-2026-LearningAgenda.pdf.

- Federal agencies and other stakeholders better understand the unique milieu of barriers within each community and use that knowledge to design integrated solutions

- Local communities find other communities like their own and identify what lessons they may be able to learn from one another that can lead to better outcomes

See Appendix Figure 3 in the appendix for an example of a MITRE-developed equitable data web dashboard.

Lesson Learned: Sharing promising practices and standards for equitable data sharing among practitioners and users bolsters the impact of equity-improving efforts.

MITRE's qualitative analysis uncovered additional data elements that mapped into the equitable data framework (see Appendix Figure 1). TTLS governments—having learned and applied strategies in connection with their federal funding allocations—can apply these data elements to address equity throughout all their operations. Further, TTLS entities can share standards on successful equity strategies and best practices to include items such as:

- Data infrastructure

- Equity assessments

- Equity evaluation methods and tools

- Goal setting

- Identifying high-need populations

- Engaging underserved communities so people can have a voice in their own future

- Collecting performance data to track progress toward advancing equitable outcomes

3. What resources, programs, training, or other tools can facilitate increased data sharing between different levels of government (Tribal, territorial, local, State, or Federal) related to equitable data?

MITRE suggests that three resources can facilitate increased data sharing among different levels of government, historically underrepresented scholars and research institutions, community-based organizations, and the public: (1) MITRE's deliberately inclusive and stakeholder-driven set of tools and approaches supporting data-sharing partnerships in the Public-Private Partnerships Accelerator Toolkit (P3TK); (2) the MITRE Social Justice Platform's Fairness, Agency, Inclusion, and Representation (FAIR) Framework for designing research on equity data focused on the lived experiences of individuals within communities; (3) outreach, training, and program-driven adoption by TTLS of equity data and data standards embodied in federal workplace surveys.

P3TK. Because the data that can best address a given topic (e.g., how well government programs and policies serve different populations) often comes from multiple sources, MITRE sees analogies to considerations within data-sharing partnerships. Based on its experiences shaping and supporting data-sharing partnerships, MITRE created P3TK for general use in accelerating successful collaboration among the affected/interested parties and has applied P3TK in

supporting a variety of partnerships. One example is the Identity Theft Task Refund Fraud (IDTTRF) Information Sharing and Analysis Center (ISAC). The ISAC involves collaboration among federal and state agencies, tax return preparation companies, and others to proactively mitigate issues with multi-sourced data, which disproportionately affects taxpayers with lower income. Through secure data sharing, the social harm of IDTTRF is ameliorated—the IRS Commissioner has noted that "thousands of taxpayers were protected from victimization as a result of the efforts of the ISAC Partnership."[8] The P3TK resource advocates for collaboration predicated on principles of trust-building, transparency, representation, and partner-driven design and operation including shared decision making. P3TK also offers practical guidance to address interagency and cross-sector collaboration; governance and protocol for collaboration; ethics, principles, agreements, and legal compliance; and other topics relevant to the affected entities (partners) actively co-designing and cooperating on their data-sharing initiative.

FAIR. While innovation and collaboration can unlock the benefit of data sharing, including community voices is crucial for improving data integrity and producing equitable, sustainable solutions (see the lesson learned in Question 2, above). When local communities share their lived experiences—knowledge, beliefs, patterns, and expectations encapsulated within daily living and cultural norms that go beyond pure quantitative statistical data—that drives better understanding of the causal factors and outcomes that should be the focus of the data-centric research and/or collaboration to improve government program design. To that end, MITRE's FAIR Framework designs equity into data sharing and analysis collaboratives through community voices and systems thinking. Since data sharing doesn't happen simply because one wants it to, the FAIR Framework provides a new way to encourage data sharing so that it is driven by stakeholders' perspectives of the problems/solutions. Without FAIR, data sharing may have limited participation/adoption and resulting models may be theoretically accurate but realistically ineffective or inequitable.[9]

The FAIR Framework is a systematic methodology that enables qualitative and quantitative research on structural factors—such as poverty, health, education, and employment—that may drive inequality. FAIR integrates community voices to identify and explore the architecture of disparities and the design of equity. This helps clarify the structural elements that lead to a community's experience of inequity or adverse outcomes such as persistent poverty or health disparities. FAIR captures factors impacting people's daily lives and identifies data and models that would have otherwise remained unanticipated and hidden to those outside of the impacted community. This approach also allows researchers to uncover unconscious biases and assumptions.

By using resources and tools like P3TK and FAIR, data sharing—across levels of government, among historically underrepresented researchers, and with community-based organizations—is foundationally accelerated and enhanced. These methods help stakeholders tangibly experience seeing themselves in, actively contributing to, and deriving benefit from the data-sharing work. This kind of practical, stakeholder-driven approach to data sharing results in substantially increased buy-in, higher levels of participation (more, and more diverse, entities actively

---

8   Security Summit. 2021. Internal Revenue Service, https://www.irs.gov/newsroom/security-summit. Last accessed September 28, 2022.

9   FAIR Framework: Designing Equity through Community Voices and Systems Thinking. 2022. MITRE, https://sjp.mitre.org/insights/60f1e225b1d934001a56df51. Last accessed September 28, 2022.

contributing to the work), and more impactful and equitable outcomes (due to the power of more data from so many perspectives, inclusion of community voices, and strengthened collaboration).

Federal survey standards. The federal government should support TTLS government adoption of federal surveys, such as the Federal Employee Viewpoint Survey[10] that measures diversity, equity, inclusion and accessibility (DEIA) and the Government Pulse Survey,[11] through investment in awareness and usage training and programs that connect adoption to grant performance and budgets. These surveys collect information on historically underrepresented groups providing further segmentation and targeting decision making and investments in more equitable ways. Increased adoption of these established questionnaires will lead to a greater set of survey results data available for federal and TTLS government, historically underrepresented scholars and research institutions, and community organizations to use. Adoption of federal survey standards will also promote clear and consistent indicators of inclusion and equity at state and local levels and related benchmarking, informing investments and interventions. Further, adopting consistent measurement approaches promotes transparency and increases trust in government practices and services.

4. What resources, programs, training, or other tools can expand opportunities for historically underrepresented scholars and research institutions to access and use equitable data across levels of government?

The examples provided for Question 3 also apply for this community. MITRE also recommends the federal government expand historically underrepresented scholars' and research institutions' access to/adoption of three data resources—Census Bureau, EEOC, and Department of Labor data—through targeted training and outreach programs. As shown by loci of researchers proximal to national laboratories, when agencies deliberately partner with entities (e.g., providing access to data, knowledge about that data, analytic and subject matter expertise, funding scholarships and grants) to foster the entity's expertise in an area, that local investment in talent and capabilities pays dividends for decades to come through workforce development, knowledge sharing (e.g., publications), innovation, and economic growth.

An issue for historically underrepresented scholars beyond access is limited knowledge and experience with equitable research practices and sources, including training on SDC data and methodologies and practices to support equity assessment for policymaking and program implementation. Educational programs that build knowledge and skills in evaluation and data analysis of TTLS data sources would be beneficial.

5. What resources, programs, training, or tools can increase opportunities for community-based organizations to use equitable data to hold government accountable to the American public?

---

[10] Governmentwide Management Reports. 2022. Office of Personnel Management, https://www.opm.gov/fevs/reports/governmentwide-reports/. Last accessed September 28, 2022.

[11] Government-Wide Pulse Survey Pilot. 2022. General Services Administration, https://d2d.gsa.gov/report/government-wide-pulse-survey-pilot. Last accessed September 28, 2022.

The examples provided for Question 3 also apply to community-based organizations.

6. What resources, programs, training, or tools can make equitable data more accessible and useable for members of the public?

The CDFI example in response to Question 2 and the examples provided for Question 3 also apply here. In addition, MITRE recommends that National Science Foundation invest in cloud environments that could be shared among the stakeholder segments (e.g., historically underrepresented researchers, the public) noted in this RFI to boost access to equitable data, and as suggested for artificial intelligence in the Open Knowledge Network Roadmap.[12]

---

[12] Open Knowledge Network Roadmap: Powering the Next Data Revolution. 2022. National Science Foundation, https://nsf-gov-resources.nsf.gov/2022-09/OKN%20Roadmap%20-%20Report_v03.pdf.
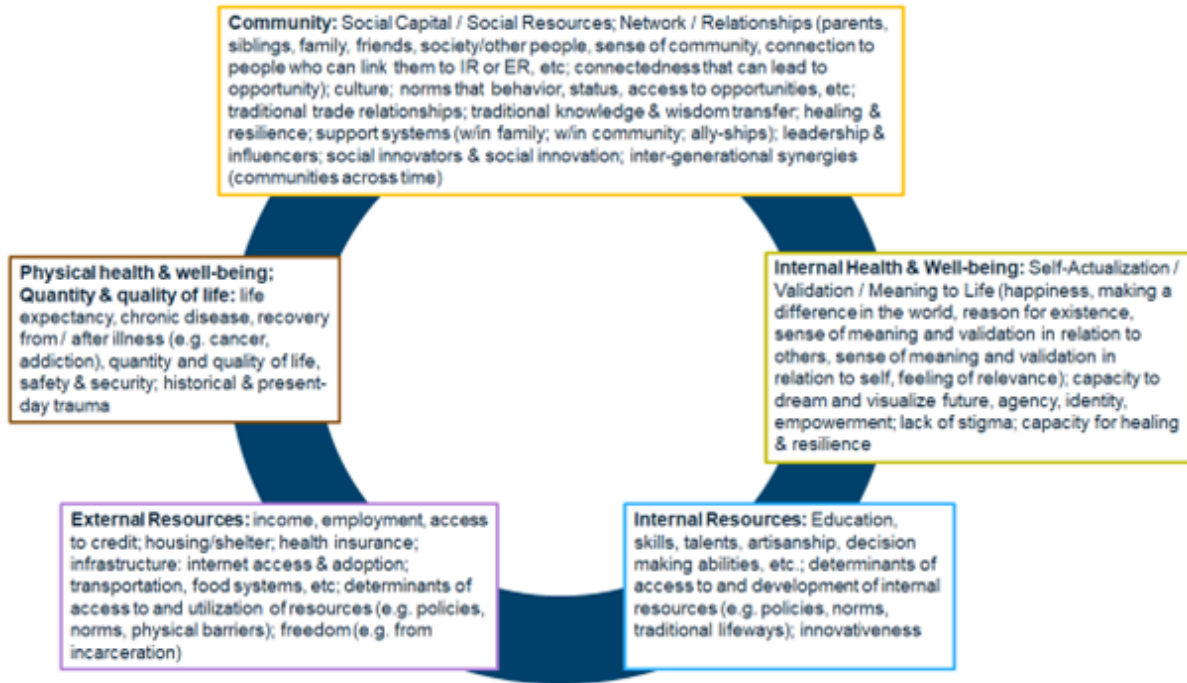
Appendix – Illustrations

Figure 1. Equitable data types derived from community interviews. Reproduced from "Voices of the Nation's Poor – A Proposed Methodology to Improve Outcomes for Persistent Poverty Counties," MITRE Technical Report MTR210506, September 2021.
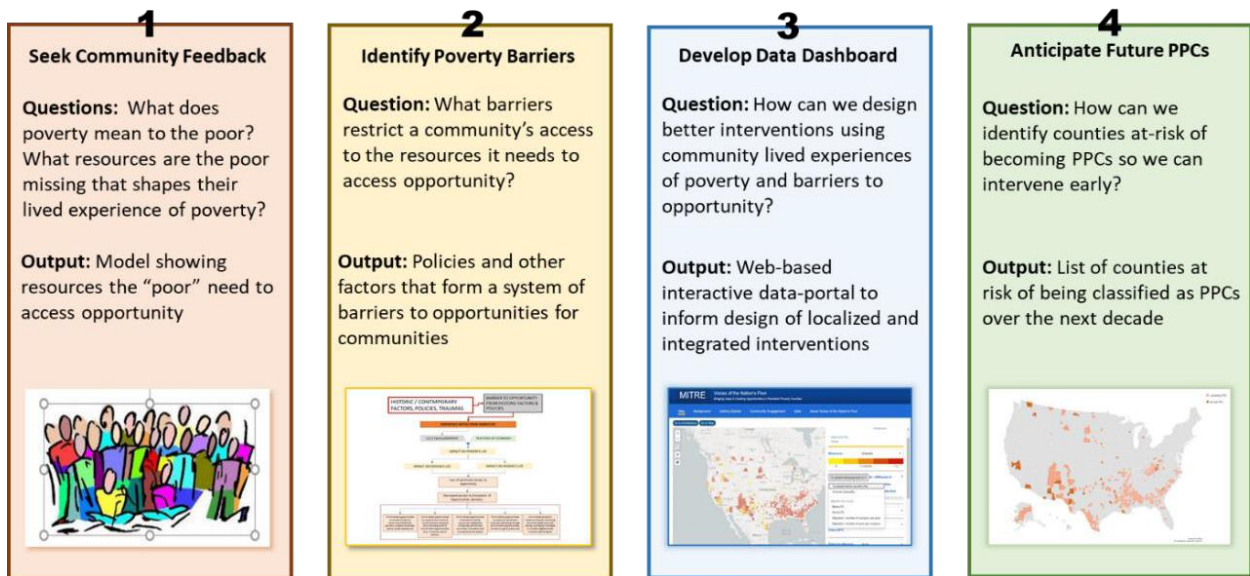

Figure 2. Four ways to expand and enrich community-centered equitable data research. Reproduced from "Voices of the Nation's Poor – A Proposed Methodology to Improve Outcomes for Persistent Poverty Counties," MITRE Technical Report MTR210506, September 2021.
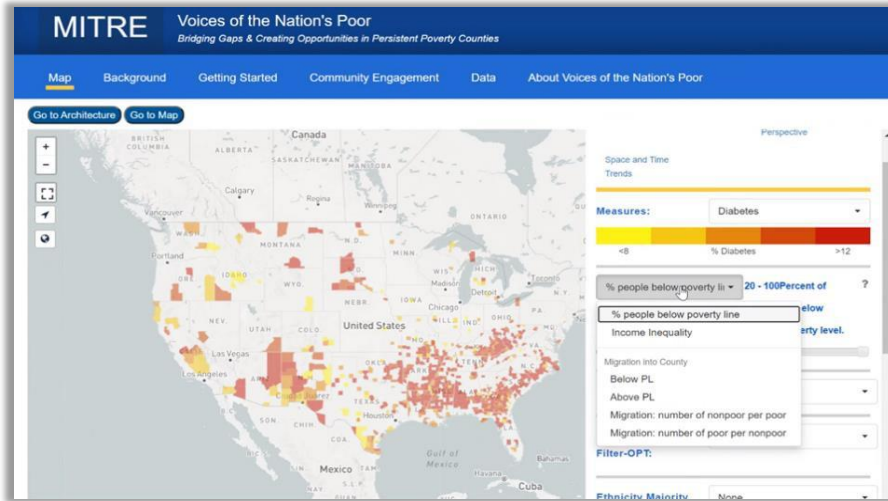
Figure 3. MITRE "Voices of the Nation's Poor" web-based equitable data dashboard. Reproduced from "Voices of the Nation's Poor – A Proposed Methodology to Improve Outcomes for Persistent Poverty Counties," MITRE Technical Report MTR210506, September 2021.